

GendCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:25:40 ; Search time 69 Seconds
(without alignments)

Title: US-09-896-580A-1
Perfect score: 1112
Sequence: 1 MYEYIANNLIFTVIQLKQIKIR.....KDHPHQPHDAVEWQHHEH 213

Scoring table: BL05UN62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Genesed 101002,*

- 1: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDS2/gcdata/genesed/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDS2/gcdata/genesed/genesed/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1112	100.0	213	AAM4343	Staphylococcus aur
2	938	87.4	183	AAW3186	Staphylococcus aur
3	817	73.5	210	ABP3601	Staphylococcus epi
4	795	71.5	183	AGB1798	S. epidermidis ope
5	754	67.8	150	AAW3187	Staphylococcus aur
6	545	49.0	146	AGB1977	S. epidermidis ope
7	535.5	48.0	183	ABP27633	Listeria monocyt
8	477	42.9	204	ABP2597	Streptococcus poly
9	464	41.7	204	ABP2596	Streptococcus poly
10	457.5	41.1	203	AAN77215	Streptococcus pneu

11	436	39.2	211	ABB53669	Lactococcus lactis
12	386.5	34.8	155	AYV86043	S. pneumoniae deri
13	386.5	34.8	155	AYV77216	Streptococcus pneu
14	220	20.7	216	AAM48347	Mycoplasma pneumo
15	211.5	19.0	169	AAM48345	Haemophilus influe
16	206	18.5	273	AGG20819	Arabidopsis thachoma
17	198	17.8	181	AAY04475	Chlamydia trachoma
18	197	17.7	158	AAG20841	Arabidopsis thalia
19	197	17.7	162	AAG20840	Arabidopsis thalia
20	186.5	16.8	160	AAY16109	A formate transpor
21	184.5	16.6	160	AAM48346	Bacillus subtilis
22	184	16.5	259	AAG28445	Arabidopsis thalia
23	184	16.5	259	AAG42460	Arabidopsis thalia
24	184	16.5	264	AAG28424	Arabidopsis thalia
25	184	16.5	269	AGA24259	Escherichia coli D
26	180.5	16.2	168	AAM48344	C. Glutamicum prote
27	180.5	16.2	169	AAM47032	Corynebacterium gl
28	179	16.1	170	AAM48348	E.coli peptide def
29	178	16.0	204	AAY36886	Staphylococcus aur
30	176	15.8	166	AAY36827	C. pneumoniae prot
31	164	14.7	403	ABP62406	Protein involved i
32	160.5	14.4	193	AAG92754	Drosophila melanog
33	160.5	14.4	193	AAG92754	C. Glutamicum prote
34	158	14.2	155	AAG22705	Corynebacterium gl
35	158	14.2	155	AAC42461	Arabidopsis thalia
36	152	13.7	136	AAB28421	Streptococcus poly
37	151.5	13.6	174	AAN96626	H. pylori GPRC 26f
38	147	13.6	163	AAB39983	Staphylococcus epi
39	143	12.9	169	AAG91517	C. Glutamicum prote
40	141.5	12.7	214	AAT9220	Propionibacterium
41	141	12.7	131	AAM62635	Streptococcus pneu
42	141	12.7	136	AAT70355	Streptococcus poly
43	140	12.6	136	AABP28422	Human mitochondria
44	140	12.6	243	AEB03548	Human peptide defo
45	140	12.6	243	AEE21799	Human peptide defo

ALIGNMENTS

RESULT 1	23-APR-2002 (first entry)
ID AAM48343	standard; protein; 213 AA.
XX	
AC AAM48343 /	
XX	
DT	
XX	
DE	Staphylococcus aureus peptide deformylase with C-terminal 6xHis tag.
XX	
KW	Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.
XX	
OS	Staphylococcus aureus.
XX	
PN WC0020258-A2.	
PD 10-JAN-2002.	
XX	
PP 29-JUN-2001; 2001IWO-US20777.	
XX	
PR 30-JUN-2000; 2000US-215550P.	
XX	
PA (PEHA) PHARMACIA & UPJOHN CO.	
XX	
PI Baldwin ET, Harris MS;	
XX	
DR WPI; 2002-146012/19.	
XX	
PT Crystalline Staphylococcus aureus peptide deformylase useful for	
PT solving structures of other molecules or molecular complexes, and	
PT designing modifiers of Peptide deformylase activity	
XX	
PS Claim 67; Fig 3; 149pp; English.	

XX The present sequence is Staphylococcus aureus peptide deformylase (pdf).
 CC The present invention relates to the crystal structure for pdf, which can
 CC be used in combination with a computer-assisted method for identifying,
 CC designing and making a potential modifier of S.aureus pdf activity.
 XX Modifiers of pdf are useful for blocking bacterial growth.

Sequence 213 AA;

Query Match 100.0%; Score 1112; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2.2e-113;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MYEVLNNLTFVILQKIKKVKYQMLTKMDIDTRGHTPLRQKAELPLTREKEETLIA 60
 Db 1 MYEVLNNLTFVILQKIKKVKYQMLTKMDIDTRGHTPLRQKAELPLTREKEETLIA 60
 Qy 61 MREFLVNSDDETAQYGLRSQVGAAPDINTSKRMATAVLIPDDSGSKSYDMLVNPKIV 120
 Db 61 MREFLVNSDDETAQYGLRSQVGAAPDINTSKRMATAVLIPDDSGSKSYDMLVNPKIV 120
 Qy 121 SHSTQEATLPTGRGCLSVDDNVAQGHVRHINKITIKAIDBENDIQRLKGPAIVFQHEI 180
 Db 121 SHSTQEATLPTGRGCLSVDDNVAQGHVRHINKITIKAIDBENDIQRLKGPAIVFQHEI 180
 Qy 181 DHLANGMFTDHDIDKDHPLQPHTDAAVEHHHH 213
 Db 181 DHLANGMFTDHDIDKDHPLQPHTDAAVEHHHH 213

RESULT 2

ID AAW83186 standard; Protein; 183 AA.
 XX AAW83186;

AC DT 11-FEB-1999. (first entry)

XX DE Staphylococcus aureus deformylase 1.
 XX Staphylococcus aureus; DefI; deformylase 1; bacterial infections;
 KW antimicrobial screening; diagnosis; antibacterial.

XX OS Staphylococcus aureus.

XX PI Lonetto MA, sylvester DR, Warren RL;

XX DR WPI; 1998-596871/51.

XX DR N-PSDB; AA70274.

XX PR 25-NOV-1998.

XX PF 21-MAY-1998; 98EP-0304076.

XX PR 15-AUG-1997; 97US-0911844.

XX PR 21-MAY-1997; 97US-0048706.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PN US6380370-B1.

XX PD 30-APR-2002.

XX XX 13-AUG-1998; 98US-0134001.

XX PF 14-AUG-1997; 97US-055779P.

XX PR 08-NOV-1997; 97US-064964P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX XX Doucette-Stamm LA, Bush D;

DR N-ISDB; AB91146.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections -
 PT PT

XX requiring reduction in Def activity, especially they are antibiotics
 CC for treating a wide range of infections caused by Staphylococcus, also
 CC by some other bacteria, e.g. Helicobacter pylori. DefI proteins are also
 CC used to screen for compounds that interact specifically with it (i.e.
 CC potential antibiotics); to study the role of Def in disease; for
 CC generating Ab and in protective vaccines (to generate an antibody and/or
 CC T-cell response). Vaccination may also be with a vector that contains a
 CC nucleic acid molecule encoding DefI. Ab are used therapeutically and for
 CC diagnosing in standard immunoassays. Antagonists may inhibit binding of
 CC bacteria to extracellular matrix proteins and to in-welling devices, or
 CC they inhibit normal progression of infection. Fragments of nucleic acid
 CC molecules encoding DefI are useful as hybridisation probes and/or
 CC amplification primers for isolation of full-length Def-encoding
 CC sequences or related genes; to detect expression in infected tissue,
 CC e.g. for diagnosis and staging, and to identify mutations (e.g. for
 CC serotyping).

XX Sequence 183 AA;

Query Match 84.4%; Score 938; DB 19; Length 183;

Best Local Similarity 98.9%; Pred. No. 2e-94; Mismatches 2; Indels 0; Gaps 0;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 MLTMKDIRDGHTPLRQKAELPLTREKEETLIAREFLNVNSODEEAKYGLRSGVG 84

Db 1 MLTMKDIRDGHTPLRQKAELPLTREKEETLIAREFLNVNSODEEAKYGLRSGVG 60

Qy 85 LAAPQINISKRMATAVLIPDDSGSKSYDMLVNPKIVSVOEAYLPTGEGLSVDDNVAG 144

Db 61 LAAPQINISKRMATAVLIPDDSGSKSYDMLVNPKIVSVOEAYLPTGEGLSVDDNVAG 120

Qy 145 MHRKAKTAKDIENDICLRKGPAIVFOHEIDHNGVMFYDHDKDHPLQPHTD 204

Db 121 LVHRNRTTAKDIENDICLRKGPAIVFOHEIDHNGVMFYDHDKDHPLQPHTD 180

Qy 205 VEV 207

Db 181 VEV 183

RESULT 3

ABP38601

ID ABP38601 standard; Protein; 210 AA.

XX ABP38601;

XX AC ABP38601;

XX DT 24-JUL-2002 (first entry)

XX DS Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3446.

XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX KW antibacterial; gene therapy.

OS OS Staphylococcus epidermidis.

XX PN US6380370-B1.

XX PD 30-APR-2002.

XX XX 13-AUG-1998; 98US-0134001.

XX PF 14-AUG-1997; 97US-055779P.

XX PR 08-NOV-1997; 97US-064964P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX XX Doucette-Stamm LA, Bush D;

DR N-ISDB; AB91146.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections -

PT PT

Disclosure: SEQ ID 3446: 267bp; English.
ABN0538 to ABN93374 represent *Staphylococcus epidermidis* open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP3514 to ABP37960. The *S. epidermidis* sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly *S. epidermidis* infections. The sequences can be used to screen for compounds able to interfere with the *S. epidermidis* life cycle or inhibit *S. epidermidis* infection.
N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site.

Sequence	210 AA;	Query Match	73.5%	Score	817;	DB-23;	Length	210;
		Best Local Similarity	74.4%	Pred. No.	4.6e-81;			
		Matches	154;	Conservative	26;	Mismatches	27;	Indels 0; Gaps 0;
		1	MATYFLNLFTVOLQKIQIKRKYQVMTMKDIDRGDTPIRKAEELPLTMECECTLLA					60
		4	I FEELMNYLCLSYLYYLIRKVTIMTKDIDRGDTPIRKAEELPLSNDKTLLA					63
		61	MREFLINSQDEEIAARYGLRSQGCLAAQOINISKMIAVILPDGSKSVDYMLNPKIV					120
		64	MREFLINSQDEEIAARYGLRSQGCLAAQINEPKRMIAVILPDGNGSVDYMLNPKIV					123
		121	SYSVQEAYLPGLPGEGLSVDNNVALVHRRNKLITKARDLEGNDIQLRKGPATVQHEI					180
		124	SYSVQEAYLPGLPGEGLSVDENIPVLVHRRVTLAQDINGDVKLRKGPATVQHEI					183
		181	DHNGTMYFDHDIDKDPHLOHTDAVEY					207
		184	DHNGTMYFDYDANEPLPKHEAVEY					210

S. epidermidis open reading frame protein sequence SEQ ID NO:690.
Straphylococcus epidermidis SR1 strain; infection; diagnosis;
vaccination; endocarditis.

WO200134809-A2. 17-MAY-2001. 09-NOV-2000; 20000MO-US330732. 09-NOV-1999; 99US-0164258. (GRAY) STAPHYLOCOCCUS EPIDERMIDIS. STAPHYLOCOCCUS EPIDERMIDIS.

Kimmerly WJ;
WPI; 2001-316495/33.
N-PSDB; AAH52648.
Nucleic acids encoding polyopeptides from *Staphylococcus epidermidis*,
useful for vaccinating against infections, e.g. endocarditis -
claim 18; Page 217; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG8145 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce host cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH51971 to AAH55050 represent specifically claimed *S. epidermidis* genomic DNA polymeric sequences from the present invention. AAG55091 to AAH55098 represent oligonucleotide sequences and primers which are used

QY	205	VEV	207	
Db	181	VEV	183	
RESULT 5				
AAW83187	AAW83187	standard;	Protein,	150 AA.
LD	XX			
AC	AAW83187;			
XX				
DT	11-FEB-1999	(first entry)		
XX				
DE	Staphylococcus aureus deformylase 1 ORF protein sequence			
XX				
KW	staphylococcus aureus; DefI; deformylase 1; bacterial in			
KW	antimicrobial screening; diagnosis; antibacterial.			
XX				
OS	Staphylococcus aureus.			

PN	ER879979-A2.
XX	
PD	25-NOV-1998.
XX	
PF	21-MAY-1998;
XX	
PR	15-AUG-1997;
PR	21-MAY-1997;
XX	
(SMIK)	SMITHKLINE BEECHAM C
PA	
XX	
PA	Lonetto MA, Sylvestre DR,
XX	

DR N-PSDB; AAV70275.
 XX PT New nucleic acid encoding polypeptide deformylase of *Staphylococcus aureus* - useful for treatment, prevention and diagnosis of bacterial infections, and for antimicrobial screening
 XX PS Disclosure; Page 21; 23pp; English.
 XX The present sequence represents deformylase I (DefI) isolated from *Staphylococcus aureus* WCU 29 (NCIMB 40771). Host cells containing vectors comprising nucleic acid molecules encoding DefI are used to produce recombinant DefI which have polypeptide deformylase (Def) activity. DefI proteins are used to treat conditions requiring Def activity, while DefI antagonists, e.g. Ab, are used to treat conditions requiring reduction in Def activity, especially when they are antibacterials for treating a wide range of infections caused by *Staphylococcus*, also by some other bacteria, e.g. *Helicobacter pylori*. DefI proteins are also used to screen for compounds that interact specifically with it (i.e. potential antibacterials), to study the role of Def in disease; for generating Ab and in protective vaccines (to generate an antibody and/or nucleic acid molecule encoding DefI). Ab are used therapeutically and for diagnosing in standard immunoassays. Antagonists may inhibit binding of bacteria to extracellular matrix proteins and to in-dwelling debris, or they inhibit normal progression of infection. Fragments of nucleic acid molecules encoding DefI are useful as hybridisation probes and/or amplification primers for isolation of full-length Def-encoding sequences or related genes, to detect expression in infected tissue, e.g. for diagnosis and staging, and to identify mutations (e.g. for serotyping).
 XX Sequence 150 AA;

Query Match 67.8%; Score 75%; DB 19; Length 150;

Best Local Similarity 98.7%; Pred. No. 2.2e-74;
 Matches 148; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 25 MLTMKDIIDGHTLQRQAELPLTKEKETLIANREFLYNSODEIAARYGLRSVG 84

Db 1 MLTMKDIIDGHTLQRQAELPLTKEKETLIANREFLYNSODEIAARYGLRSVG 60

Qy 85 LAAPQINISKRMTAVLIPDDGSGRSYDYMVLVNPKIVSHSVEAQLPTGEGLSVDDNVAQ 144

Db 61 LAAPQINISKRMTAVLIPDDGSGRSYDYMVLVNPKIVSHSVEAQLPTGEGLSVDDNVAQ 120

Qy 145 LYHRENKITKARDIEGNDICIRLGYPAM 174

Db 121 LYHRENKITKARDIEGNDICIRLGYPAM 150

Qy 145 LYHREN 150

Db 121 LYHRENH 126

RESULT 6

AAC82977 ID AAC82977 standard; Protein: 146 AA.

XX AC AAC82977;

XX DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:3048.

XX KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis;

KW vaccination; endocarditis.

XX OS *Staphylococcus epidermidis*.

XX PN WO200134809-A2.

XX PD 17-MAY-2001.

XX PF 09-NOV-2000; 2000WO-US0788.

XX PR 09-NOV-1999; 99US-0164258.

XX AF

SEQ ID NO:4455 to 4464.

XX

Query Match 49.0%; Score 545; DB 22; Length 146;
 Best Local Similarity 81.7%; Pred. No. 1.7e-51;
 Matches 103; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Qy 25 MLTMKDIIDGHTLQRQAELPLTKEKETLIANREFLYNSODEIAARYGLRSVG 84

Db 1 MLTMKDIIDGHTLQRQAELPLTKEKETLIANREFLYNSODEIAARYGLRSVG 60

Qy 85 LAAPQINISKRMTAVLIPDDGSGRSYDYMVLVNPKIVSHSVEAQLPTGEGLSVDDNVAQ 144

Db 61 LAAPQINISKRMTAVLIPDDGSGRSYDYMVLVNPKIVSHSVEAQLPTGEGLSVDDNVAQ 120

Qy 145 LYHREN 150

Db 121 LYHRENH 126

RESULT 7

ABB47633 ID ABB47633 standard; Protein: 183 AA.

XX AC ABB47633;

XX DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #337.

XX KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

XX OS Vitamin B12; bacterial infection; disease.

XX PN WO200177335-A2.

XX PD 18-OCT-2001.

XX PR 11-APR-2001; 2001WO-FR01118.

XX AF

Db 126 EGGCLSYDRNPGVVRHARVTVDYFDGGERHLKLYNSIVQHEIDHNGIMFYD 185
 Qy 191 HIDKDHP 197
 Db 186 RINEKDP 192

RESULT 11
 ID ABB53869 standard; Protein; 211 AA.
 XX
 AC ABB53869;
 XX 16-MAY-2002 (first entry)
 XX DE Lactococcus lactis protein def.
 XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX OS Lactococcus lactis IIL1403.
 XX PN FR2807446-A1.
 XX PD 12-OCT-2001.
 XX PP 11-APR-2000; 2000FR-0004630.
 XX PR 11-APR-2000; 2000FR-0004630.
 XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 DR WPI; 2002-043118/06.
 XX PT New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species -
 XX PS Claim 6; SEQ ID No 571; 2504PP; French.
 XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (AB490521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: the sequence data for this patent is based on equivalent Patent
 CC WO20017734 (published 18 OCT 2001) which is available in electronic
 CC format directly from WIPO ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 211 AA;

Query Match 39.2%; Score 436; DB 23; Length 211;
 Best Local Similarity 47.9%; Pred. No. 2.7e-39;
 Matches 93; Conservative 38; Mismatches 51; Indels 12; Gaps 3;

Qy 16 QIKRKVQYMLMKDIDRGHPTIQRKAELPLTKEKEKTLIAMREFLWNSQDEEIAK 75
 Db 7 QANLTKASMSMDDIREGPTIREVANDVPLPSDEDILGERMLOELHNSQDFVMAE 66

Qy 76 RYGHRSCLGAAPOINISKRMIAVLP----DGS---GKSYDY--MLNPKIVSHS 123
 Db 67 RQGRGQVGLAANQLGKLYKAVLIPGECKLPSYDDAVAGVHRENKTTIKAKDLEGNDIQLRLKGYPATIVQHEIDHL 126

Qy 124 VQEAVLPSEGCLSYDRNPGVVRHARVTVDYFDGGERHLKLYNSIVQHEIDHNGIMFYD 183
 Db 127 VODAAVEGEGCLSYDRNPGVVRHARVTVDYFDGGERHLKLYNSIVQHEIDHNGIMFYD 186

Qy 184 NGVMFYDHIDKDP 197
 Db 187 NGVMFYDHINMNDP 200

RESULT 12
 ID AAY86043 standard; Protein; 155 AA.
 XX
 AC AAY86043;
 XX DT 10-APR-2000 (first entry)
 XX DE S. pneumoniae derived protein #252.
 XX KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
 XX bacterial; antimicrobial; antibiotic; pathogenesis; infection.
 XX OS Streptococcus pneumoniae.
 XX PN W09806734-A1.
 XX PD 19-FEB-1998.
 XX PF 15-AUG-1997; 97WO-JS14436.
 XX PR 16-AUG-1996; 96US-0024022.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Black MT, Hodgson JB, Knowles DJC, Loretto MA, Nicholas RC;
 XX PR Scodola RR;
 XX DR WPI; 1998-159452/14.
 XX PT Streptococcus pneumoniae proteins and related DNA - useful for
 PT screening compounds for antibacterial activity
 XX PS Claim 5; Page 528; 640pp; English.
 XX DR N-PSDB; AAZ96374.

This invention describes novel isolated Streptococcus pneumoniae
 CC polypeptides (see AAZ96173-Z96494) and their encoded proteins (see
 CC AAY85792-R86182). The DNA, vectors and host cells described in the
 CC method of the invention are useful for the recombinant expression of the
 CC polypeptides. The polypeptides are useful for treatment or prevention of
 CC disease, or diagnosis of disease related to expression or activity of
 CC such a polypeptide. They can also be used to screen for compounds which
 CC interact with and inhibit or activate such a polypeptide. The
 CC polypeptides (or DNA encoding them, via Gene therapy) are also useful
 CC for inducing an immunological response in a mammal. The antagonists are
 CC useful to inhibit such bacterial polypeptides. The polypeptides are
 CC particularly useful to identify antimicrobial compounds and antibiotics.
 CC They are also useful to determine their role in pathogenesis of
 CC infection, dysfunction and disease.

Sequence 155 AA;

Query Match 34.8%; Score 386.5; DB 19; Length 155;
 Best Local Similarity 54.2%; Pred. No. 4.4e-34;
 Matches 28; Mismatches 31; Indels 7; Gaps 2;

Qy 61 MREFLVNSQDEEIAKRYGLRGSVGLAAPOINISKRMIAVLP----DGS---GKSYDY--MLNPKIVSHS 113
 Db 1 MMQFLKHSDPDMWAERGQVGLAAPOLDLSKRITIAVLPNIVEGETPQEAQYDLEAI 60

Qy 114 LVNPKVSSVQDYLPTGECKLPSYDDAVAGVHRENKTTIKAKDLEGNDIQLRLKGYPATIVQHEIDHL 173
 Db 61 MYNPKVSSVQDYLPTGECKLPSYDDAVAGVHRENKTTIKAKDLEGNDIQLRLKGYPATIVQHEIDHL 120

Qy 174 IVFOEIDHNGNMFYDHIDKDP 197
 Db 121 IVVQHEIDHNGIMFYDTRINEKDP 144

AAW77216 standard; Protein; 155 AA.
 ID AAW77216;
 AC AAW77216;
 XX XX 16-NOV-1998 (first entry)
 DE Streptococcus pneumoniae def1 open reading frame peptide sequence.
 KW Streptococcus pneumoniae; otitis media; conjunctivitis; meningitis;
 KW pneumonia; bacteraemia; sinusitis; pleural empyema; endocarditis; Def1.
 OS Streptococcus pneumoniae.
 XX EP863152-A2.
 XX PR 09-SEP-1998.
 XX PF 05-FEB-1998; 9BEP-0300873.
 XX PR 15-DEC-1997; 97HS-0311503.
 XX PR 10-FEB-1997; 97US-037536.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX PI Black MT, Hodgson JE, Knowles DSC, Lonetto MA, Nicholas RO,
 Scodola RK;
 XX WPI: 1998-458798/40.
 DR N-PSDB; AAV48309.
 XX PT New DNA encoding pneumococcal deformylase polypeptides - used to
 prevent, treat, diagnose and vaccinate against e.g. otitis media,
 conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis,
 pleural empyema and endocarditis.
 XX PS Claim 23; Page 5; 25pp; English.
 XX The Streptococcus pneumoniae Def1 protein can be used in the treatment,
 prevention and diagnosis of and vaccination against otitis media,
 conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis, pleural
 empyema and endocarditis.
 XX Sequence 155 AA;
 Query Match 34.8%; Score 386.5; DB 19; Length 155;
 Best Local Similarity 54.2%; Pred. No. 4.4e-34;
 Matches 78; Conservative 28; Mismatches 31; Indels 7; Gaps 2;
 QY 61 MREFLYNSQDEIAKRYLRSRGVGLAAPQINISKRMIAVLIPD----DGSGRSYD-M 113
 Db 1 MMCFLKHSQDPVMAEKNGLRRGVGLAAPQINISKRMIAVLIPNIVEGETPQAYDLEI 60
 QY 114 LVNPKIVSHSVQRAYLPGEGLSYDNTVAGLYRHEKKTIKAQLEGNDIQLRLKGYPA 173
 Db 61 MYNPKIVSHSVQRAYLPGEGLSYDNTVAGLYRHEKKTIKAQLEGNDIQLRLKGYPA 120
 QY 174 IYFQHEIDHNGMYFDHDKDHP 197
 Db 121 IYVQHEIDHNGMYFDRINEKDP 144
 RESULT 14
 AAM48347
 ID AAM48347 standard; protein; 216 AA.
 AC AAM48347;
 XX XX 23-APR-2002 (first entry)
 DE Mycoplasma pneumoniae peptide deformylase.
 KW Mycoplasma pneumoniae peptide deformylase.

KW Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.
 XX OS Mycoplasma pneumoniae.
 XX PN WO200202758-A2.
 XX DT 10-JAN-2002.
 XX PP 29-JUN-2001; 2001WA-US00777.
 XX PR 30-JUN-2000; 2000US-215550P.
 PA (PHAA) PHARMACTA & UPJOHN CO.
 XX PI Baldwin ET, Harris MS;
 XX DR WPI; 2002-148012/19.
 XX PT Crystalline Staphylococcus aureus peptide deformylase useful for
 solving structures of other molecules or molecular complexes, and
 designing modifiers of peptide deformylase activity -
 XX PS Example 1; Fig 3; 149pp; English.
 CC The present invention relates to the crystal structure for Staphylococcus
 aureus peptide deformylase (pdf; see AAM4343 for pdf protein sequence),
 which can be used in combination with a computer-assisted method for
 identifying, designing and making a potential modifier of S. aureus pdf
 activity. Modifiers of pdf are useful for blocking bacterial growth. The
 present sequence is Mycoplasma pneumoniae peptide deformylase (pdf),
 which was used in a sequence alignment with the S. aureus pdf sequence.
 XX Sequence 216 AA;
 Query Match 20.7%; Score 230; DB 23; Length 216;
 Best Local Similarity 30.4%; Pred. No. 1e-16;
 Matches 59; Conservative 44; Mismatches 83; Indels 8; Gaps 5;
 PS QY 7 NUFTVIQ-LKQIKTIRKQYMLTMKDLIIRDIGHPTUROKABAEFLPLITKEKEETIJAMREFL 65
 Db 11 SFRTRILILQI---NMEILLPTKAWIYLDVKELNEPTPKVDFDQASIDCRAKMWAY 67
 QY 66 VNSQDEIAKRYLRSRGVGLAAPQINISKRMIAVLIPDGGSKSDYMLVNPKIVSHSYQ 125
 Db 68 DASYNED-AKRYGUTIPGIGIGIANGQGYKQMFYTHLD--GOVERKCLLNPKTNLNSAN 124
 QY 126 EAYLPLGEGLSYDNTVAGLYRHEKKTIKAQLEGNDIQLRLKGYPA 184
 Db 125 KSFLLKGEGCLSYSPKMGQGYVIRHEWITITGDFWLQKETITATSLFGMCLEHFDLHQ 184
 QY 185 GMYFDHIDKDHP 198
 Db 185 GRAYYHRINPLNPL 198
 RESULT 15
 AAM48345
 ID AAM48345 standard; protein; 169 AA.
 XX AC AAM48345;
 XX DT 23-APR-2002 (first entry)
 DE Haemophilus influenzae peptide deformylase.
 KW Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.
 AC OS
 XX OS Haemophilus influenzae.
 PN WO200202758-A2.
 XX PD 10-JAN-2002.
 XX XX

PF 29-JUN-2001; 2001WO-US20777.
 XX
 PR 30-JUN-2000; 2000US 215550P.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Baldwin ET, Harris MS;
 XX
 DR WPI; 2002-148012/19.
 XX
 PT Crystalline Staphylococcus aureus peptide deformylase useful for
 solving structures of other molecules or molecular complexes, and
 designing modifiers of peptide deformylase activity -
 XX
 PS Example 1; Fig 3; 149pp; English.
 XX
 CC The present invention relates to the crystal structure for Staphylococcus
 CC aureus peptide deformylase (pdf; see PAM4343 for pdf protein sequence),
 CC which can be used in combination with a computer-assisted method for
 CC identifying, designing and making a potential modifier of S.aureus pdf.
 CC Activity. Modifiers of pdf are useful for blocking bacterial growth. The
 CC present sequence is Haemophilus influenzae peptide deformylase (pdf),
 CC which was used in a sequence alignment with the S. aureus pdf sequence.
 XX
 SQ Sequence 169 AA;
 Qy Query Match 19.0%; Score 211.5; DB 23; Length 169;
 Best Local Similarity 35.1%; Pred. No. 7.7e-15;
 Matches 53; Conservative 31; Mismatches 48; Indels 19; Gaps 6;
 Db Database 49 PLTKREKEKELIAMEFLVNSODEEIAKRYGLRSQVGAAPQINISKRMIAVLIPDGSGK 108
 50 PVTKND---ARKIVDMDPTMYQEK----GIGAAPOVDILQRITIDFEGD---K 67
 51 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 52 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 53 QNQFVLINEILAA--SGETGLESCLSL-PGFRALVPRCEKVTYRALDRDSKEFTLDA 123
 Qy 169 KGYPAVLFQHEIDHLNGMFTDHIDDKHPPHQ 199
 Db 124 DGLIAICIQHEIDHLNLGVFDYLSS--PLK 151

Search completed: July 11, 2003 09:35:40
 Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: July 11, 2003, 09:34:30 ; search time 26 Seconds

(without alignments)

241.042 Million cell updates/sec

Title: US-09-896-580A-1

Perfect score: 1112

Sequence: 1 MYEKLNNLFTVQLKQIKIR.....KDHPLQHFTDAVEVHQHHHH 213

Scoring table: BLOSUM62

Gapext 0.5

Searched: 262574 seqs., 29422922 residues

Total number of hits satisfying chosen parameters:

262574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:
 1: /cgn2_6/podata/1/iaa/5A_COMB.dep:
 2: /cgn2_6/podata/1/iaa/5B_COMB.dep:
 3: /cgn2_6/podata/1/iaa/6A_COMB.dep:
 4: /cgn2_6/podata/1/iaa/6B_COMB.dep:
 5: /cgn2_6/podata/1/iaa/FCITS_COMB.dep:
 6: /cgn2_6/podata/1/iaa/bachtlesi.dep:
 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	938	84.4	183	4	US-09-373-953-2	Sequence 2, Appli
2	817	73.5	210	4	US-09-134-001C-3446	Sequence 3446, Appli
3	754	67.8	150	4	US-09-373-953-4	Sequence 4, Appli
4	457.5	41.1	203	4	US-08-991-023-2	Sequence 2, Appli
5	386.5	34.8	155	4	US-08-991-023-4	Sequence 4, Appli
6	198	17.8	181	2	US-08-932-142-2	Sequence 2, Appli
7	198	17.8	181	4	US-09-342-458-2	Sequence 2, Appli
8	186.5	16.8	160	4	US-09-194-146-8	Sequence 8, Appli
9	180.5	16.2	169	2	US-08-895-939-4	Sequence 4, Appli
10	180.5	16.2	169	4	US-09-188-820-4	Sequence 4, Appli
11	147	13.2	163	4	US-09-134-001C-4828	Sequence 4828, Appli
12	89	8.0	689	1	US-08-469-005A-10	Sequence 2, Appli
13	89	7.6	2509	1	US-09-721-832-2	Sequence 10, Appli
14	84	7.6	2511	4	US-09-261-907-2	Sequence 2, Appli
15	81	7.3	638	2	US-08-557-122A-38	Sequence 38, Appli
16	81	7.3	638	4	US-09-262-666-38	Sequence 38, Appli
17	81	7.3	1447	4	US-09-376-330-17	Sequence 17, Appli
18	7.1	1375	4	US-09-122-139-2	Sequence 2, Appli	
19	79.5	7.1	1375	4	US-09-721-832-2	Sequence 2, Appli
20	79.5	7.1	1375	4	US-09-721-689-2	Sequence 2, Appli
21	78.5	7.1	600	3	US-09-212-971-12	Sequence 12, Appli
22	78.5	7.1	600	4	US-08-800-929A-12	Sequence 12, Appli
23	78.5	7.1	600	4	US-09-617-05A-12	Sequence 12, Appli
24	78.5	7.1	604	1	US-08-487-753-2	Sequence 2, Appli
25	78.5	7.1	604	2	US-08-480-065-2	Sequence 2, Appli
26	78.5	7.1	604	3	US-08-497-744-2	Sequence 2, Appli
27	78.5	7.1	604	5	PCT-US93-09167-2	Sequence 2, Appli

RESULT 2

US-09-373-953-2

Query Match Score 938; DB 4; Length 183;
 Best Local Similarity 98.9%; Freq. No. 1.2e-96;
 Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 25 MLTFMDIIRDGHPTLROAAELPLTKEETLIAREFVNSDDEFAKRGLRSVG 84

Db 1 MLTFMDIIRDGHPTLROAAELPLTKEETLIAREFVNSDDEFAKRGLRSVG 60

Qy 85 LAAPQINISKRMIAVILPDDGSKSYDYMIVNPKVSHSYOBAYLPTGEGCLSVYDDNYAG 144

Db 61 LAAPQINISKRMIAVILPDDGSKSYDYMIVNPKVSHSYOBAYLPTGEGCLSVYDDNYAG 120

Qy 145 LVHRANKITKADIEGNDIDQLRKKGYPAPIVQHEIDHLNGMVFHDIDKDHFLQPTDA 204

Db 121 LVHRANRITKADIEGNDIDQLRKKGYPAPIVQHEIDHLNGMVFHDIDKDHFLQPTDA 180

Qy 205 VEV 207

Db 181 VEV 183

RESULT 2

US-09-134-001C-3446

Sequence 3446, Application US/09134001C

; Patent No. 6,803,70

; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIORITY NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 SEQ ID NO: 3446
 LENGTH: 210
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3446

Query Match Score 817; DB 4; Length 210;
 Best Local Similarity 74.4%; Pred. No. 4.4e-83;
 Matches 154; Conservative 26; Mismatches 27; Indels 0; Gaps 0;

Qy 1 MYEYLNNLFLTVIQLQKIKRKYQVMTMDIIRDGHPTLRQRAAELLTKEKEKTLIA 60
 Db 4 IFEYLLMLCLSLVLYLIRKVDITMMDIIRDGHPTLRKAEKLSTPISMDKETLRA 63

Qy 61 MREFLINNSODEEIAKRYGILRSVGYLAAPQINISKRMIAVLIPDDGSKSYDMLVNPKIV 120
 Db 64 MREFLINNSODEEIAKRYGILRSVGYLAAPQINEPGRMIAVLFDGGSKSYDMLVNPKIM 123

Qy 121 SHSVQAYLFTGEGLSYDNDVAGLVHRINKITKADEGNDIQLRKGYPAVTFQHEI 180
 Db 124 SYSVQAYLPTGEGLSYDNTPLGVHRHVTIKAQDGDVNRKURKGYPAAVFEQHEI 183

Qy 181 DHLNGMFYDHDQPHDQPHDAVEY 207
 Db 184 DHLNGMFYDIDANEPLKPHBEEAVY 210

RESULT 3
 US-09-373-953-4
 Sequence 4, Application US/09373953
 Patent No. 6410688
 GENERAL INFORMATION:
 APPLICANT: Loretto, Michael A.
 APPLICANT: Sylvester, Richard
 APPLICANT: Warren, Richard
 TITLE OF INVENTION: No. 6410688el def1
 FILE REFERENCE: GM00001
 CURRENT APPLICATION NUMBER: US/09/373,953
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 08/911,844
 PRIOR FILING DATE: 1997-08-15
 PRIOR APPLICATION NUMBER: US 60/048,706
 PRIOR FILING DATE: 1997-05-21
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 4
 LENGTH: 150
 TYPE: PRT
 ORGANISM: Staphylococcus aureus
 US-09-373-953-4

Query Match Score 67.8%; Score 754; DB 4; Length 150;
 Best Local Similarity 98.7%; Pred. No. 2.7e-46;
 Matches 148; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 25 MLTMKDIDRGHPTLRQRAAELPLTKEKEKTLIAFRLVNSODEEIAKRYGLRSVG 84
 Db 1 MLTMKDIDRGHPTLRQRAAELPLTKEKEKTLIAFRLVNSODEEIAKRYGLRSVG 60

Qy 85 LAAPQINISKRMIAVLIPDDGSKSYDMLVNPKIVSHSYQEAYLPTGEGLSYDNTAG 144
 Db 61 LAAPQINISKRMIAVLIPDDGSKSYDMLVNPKIVSHSYQEAYLPTGEGLSYDNTAG 120

RESULT 4
 US-08-991-023-2
 Sequence 2, Application US/08991023
 Patent No. 6284978
 GENERAL INFORMATION:
 APPLICANT: Loretto, Michael A.
 TITLE OF INVENTION: def1
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Decher Price & Rhoads
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Strre
 CITY: Philadelphia
 STATE: PA
 COUNTRY: US
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/991,023
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Falk, Stephen T
 REGISTRATION NUMBER: 36,795
 REFERENCE/DOCKET NUMBER: GM50010
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-594-2488
 TELEX: 215-594-2222
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 203 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-991-023-2

Query Match Score 457.5%; DB 4; Length 203;
 Best Local Similarity 50.3%; Pred. No. 4.7e-43;
 Matches 94; Conservative 37; Mismatches 49; Indels 7; Gaps 2;

Qy 18 KIRKVQMLMKTIDRGHPTLRQRAAELPLTKEKEKTLIAFRLVNSODEEIAKRY 77
 Db 6 RITKAHLIIMDIDIREGNPKIRTYAETTFLSDQETILSERKOMQFLHSQDFVMAEKW 65

Qy 78 GERSGYGLAPQINISKRMIAVLIPDDGSKSYDNTAG - - - DSGSKSYD - - -
 Db 66 GURGGYGLAPQLDISKRIAVLNIVEEGTPQEADLZEMIMNPKVSHSYQDAGL 125

Qy 131 TGERGLSYDNDVAGUHHRNKTIAKDEGNDIQLRKGYPAVFEQHEDHANGMFYD 190
 Db 126 EGPGLSYDNDVNPQSYVWRHARVTDYFDKGKERIKLKGYNSIVQHELDHANGMFYD 185

Qy 191 HIDKDHP 197
 Db 186 RTNEKOP 192

RESULT 5
 US-08-991-023-4
 Sequence 4, Application US/08991023

Patent No. 6284878
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: def1
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q.
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-932-142-2
; Query Match Score 17.8%; DB 2; Length 181;
; Best Local Similarity 33.5%; Pred. No. 3.2e-14;
; Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;
; Qy 31 IIRD---GHTPIRQKAEELPLTKEKEETIJJAREFLYNSQDEEIAKRYGRSGYGLA 86
; Db 1 MIRDLEYDSPILRKVAAPV TEITDELRQLLDMSETM-----AFYKGYGLA 47
; Qy 87 APOINTSKRMAYLIP--DGSGRSKSYD--MLVNPKIVSHSVOEAYLPTGECLSYDDN 141
; Db 48 APQVQSIISLFINGVERELENGLYCDEPRFFINPVTOKSEQLVY-GNECLS1-LR 104
; Qy 142 VAGLYFRHMKITKAIDEGNDIQLRKSYPAVFQEBIDHNGMFYDH1-DKDHLQ 199
; Db 105 IRGETVARPDKTESAKNLDGQSLAEGFLARIVMEHTDEHLGYLYDRMSDXDKTKQ 163
; RESULT 7
; US-09-942-458-2
; Sequence 2, Application US/09342458
; Patent No. 637022
; GENERAL INFORMATION:
; COMPUTER: IBM Compatible
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: NOVEL def
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; Sequence 2, Application US/08932142
; Patent No. 592666
; GENERAL INFORMATION:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/342,458
; FILING DATE: 29-Jun-1999
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,142
; FILING DATE: 16-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q.
; REGISTRATION NUMBER: 28,354

REFERENCE/DOCKET NUMBER: GM10095
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-994-2222
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 181 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SBQ ID NO: 2:
 US-09-342-458-2

Query Match 17.8% Score 198; DB 4; Length 181;
 Best Local Similarity 33.5%; Pred. No. 3.2e-14;
 Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;

Qy 31 IIRD---GHPTRQKAELBLPLTKEKEKETLIAAMREPLVNSQDEETIAKYGLRSGYGLA 86
 Db 1 MIRDLEYDPSPLRKVAPV-TEITIDRLPMLNSETM-----ATYKGVGGLA 47

Qy 87 APQINISKERNIAVILIP--DDGSGSKSYD--MLYNPKIVSHSVQRAYLPTGEGCLSYDDN 141
 Db 48 AFQVGQSISIFIMVSVEREDELVCFDFRVEFNPVTKRSQEQLV--GNEGGLSL-LR 104

Qy 142 VAGLVRHNRNKKITIAKEDFGCNQYPLKQYPAVHQHIDHNGMPYDH-DYDAFLQ 199
 Db 105 IRGEVARPDKITESAKNQDGQOLSLALEBPLARIVMHTDHLHGVLYIDRMSDDKTKQ 163

RESULT 8
 US-09-194-146-8
 Sequence 8, Application US/09194146
 GENERAL INFORMATION:
 APPLICANT: Miller, Brian
 ATTORNEY: Diaz-Torres, Maria
 TITLE OF INVENTION: Gram-Positive Microorganism Formate Pathway
 FILE REFERENCE: GC35-US
 CURRENT APPLICATION NUMBER: US/09/194,146
 CURRENT FILING DATE: 1998-11-19
 PRIOR APPLICATION NUMBER: GB 9724627.6
 PRIOR FILING DATE: 1997-11-20
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FastSBQ for Windows Version 3.0
 SEQ ID NO: 8
 LENGTH: 160
 TYPE: PRT
 ORGANISM: Bacillus

Qy 26 LTMKDIRDGHPT-IROGAELBLPLTKEKEKETLIAAMREPLVNSQDEETIAKYGLRSYVG 84
 Db 1 LAVKKT-T-HPAEVLTPAETVTFDKLKLDDMTMVE-----MDSVG 46

Qy 85 LAAPQINISKERNIAVILPDDGSGSKSYDMLYNPKIVSHSVQRAYLPTG-RGCLSVDDNA 143
 Db 47 LAAPQITGILKRAAEVIGGD-RER---IDLVNPELEKSGEQ---TGIGECLSF-PRVY 97

Qy 144 GLYVRHNRNKKITIAKEDFGCNQYPLKQYPAVHQHIDHNGMPYDH-DYDAFLQ 194
 Db 98 GDYTRADYYKVAFNRCRKPFLEARQHEMHDLYGKLFMDYLSS--PLK 151

RESULT 10
 US-09-188-820-4
 Sequence 4, Application US/09188820
 Patent No. 6196902
 GENERAL INFORMATION:
 APPLICANT: Bogosian, Gregg
 TITLE OF INVENTION: Deformylation of f-Met Peptides in Bacterial Expression Systems
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: U.S.
 ZIP: 77210
 COMPUTER READABLE FORM:

RESULT 9
 US-08-895-939-4
 Sequence 4, Application US/08895939
 Patent No. 5834243

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/188,820
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/895,939
 ATTORNEY/AGENT INFORMATION:
 NAME: Patterson, Mellinda L.
 REFERENCE/DOCKET NUMBER: 33-062
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (713) 787-1400
 TELEX/FAX: (713) 789-2679
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 169 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-188-820-4

Query Match 16.2%; Score 180.5; DB 4; Length 169;
 Best Local Similarity 32.9%; Pred. No. 2.6e-12;
 Matches 53; Conservative 31; Mismatches 54; Indels 23; Gaps 8;

QY 46 LEPLITKEKEKLIAMREFLYNSQ----DEEIAKEYGLESNGVLAAPQINISKRMIAVL 100
 Db 7 LHIP---DRLRKVKAVPVEIYQVNRNQVTKTAKD 158

QY 101 IPDGSGSKSYDMLVNPKVSHSYQEAFLPTG--ESCLSTYDDNVAGLYHFRNKITAKD 158
 Db 62 --DYSENDERIVLNLPEKSEB----TQEESCLSSPEQRA-LVPRAEVKRALD 113

QY 159 IEGNDIQRLRKYCPAIVFQHEIDHLANGWYFDHDKDHPIQ 199

Db 114 RDGKPPFELEADGLIAICIQHEMDHLVGKLFMDYLs--PRK 151

RESULT 11
 US-09-134-01C-4828
 Sequence 4828, Application US/09134001C
 GENERAL INFORMATION:
 Patent No. 6380370
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GT01-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO: 4828
 LENGTH: 163
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis

Query Match 13.2%; Score 147; DB 4; Length 163;
 Best Local Similarity 25.3%; Pred. No. 1.3e-08;
 Matches 46; Conservative 41; Mismatches 71; Indels 24; Gaps 7;

QY 25 MLTKKDIRDGHTPLRQKAELSLPTEKEETLAMRELVNSQDETTAARYGRSGNG 84
 Db 1 VMTVKLWKSTRPLINKTIQ---PVSTYDQKLKVLEDLTUHEAA-----A 47

QY 85 LAAPQINISKRMIAVLIPDCGSKSYDMLVNPKVSHSYQEAFLPTGEGCLSVDDNVAG 144
 Db 48 ISAPJIGVDQSTALIMEQEE---IQLINFVVGSQS-QET-YSDELGSISL-PHYG 99
 QY 145 LYHRHKNTIKTAKDIEGNDIQRLKGYPATIVFOHEIDHLANGMF--YDHDKDHPHQH 201
 Db 100 EVKRSKMTIVSQSTDINGNAVELTAYDIARMTHMIDHLANGQFTKRAHHILNETEVEAY 159

QY 202 TD 203
 Db 160 FD 161

RESULT 12
 US-08-248-021A-2
 Sequence 2, Application US/08248021A
 Patent No. 5648240
 GENERAL INFORMATION:
 APPLICANT: Hook, Magnus
 APPLICANT: Jonsson, Klas
 APPLICANT: Patti, Joseph M.
 APPLICANT: Gurusiddappa, Sivarshankarappa
 TITLE OF INVENTION: MHC II ANALOG FROM STAPHYLOCOCCUS AUREUS
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESS: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: United States of America
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/248,021A
 FILING DATE: 24-MAY-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Kitchell, Barbara S.
 REGISTRATION NUMBER: 33,928
 REFERENCE/DOCKET NUMBER: TAKM:155
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (512) 474-7577
 SEQUENCE CHARACTERISTICS:
 LENGTH: 689 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-248-021A-2

Query Match 8.0%; Score 89; DB 1; Length 689;
 Best Local Similarity 21.2%; Pred. No. 0.35%;
 Matches 36; Conservative 27; Mismatches 63; Indels 44; Gaps 4;

QY 4 YLNNEFTVQLRQIKRKVQMLTMKDILRDGHTPLTKEKTLAMRE 63
 Db 557 FTRNLPSVKDIIKIDINVQOKSKENKALK---VTINKATKVKEPVTINGFSNLVS-NE 610

QY 64 FL-----VNQDDEEAKRYGLRSGVLAAPQINISKRMIAVLIPDDGSGKSYDMLVNP 117
 Db 611 FAFLPHHTITNDLNKLRLALRSDQGITKHIDIGLERITYKVFQDSSKLEDLKAA-- 668

QY 118 KIVSHSVQEAFLPTGEGCLSYDNTAGLYHFRNKTPKAYDIEGNDIQLR 167
 Db 669 -----KQDSKVFKATDILKVDIEK 688

US-08-469-05A-10
i Sequence 10, Application US/08469005A
i Patent No. 565874
; GENERAL INFORMATION:
; APPLICANT: KUHADTA, FRANCIS P.
; APPLICANT: PASTERNAK, GARY A.
; TITLE OF INVENTION: CANCER RELATED ANTIGEN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEES: BAKER & BORTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,005A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/188,426
; FILING DATE: 24-JAN-1994
; APPLICATION NUMBER: 08/096,908
; FILING DATE: 26-JUL-1993
; APPLICATION NUMBER: 07/917,716
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 062482-0113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2509 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-469-05A-10

Query Match 7.6%; Score 84; DB 1; Length 2509
Best Local Similarity 21.6%; Pred. No. 9.3;
Matches 27; Conservative 32; Mismatches 40; Indels 26; Gaps 3;
Query 5 INNLFFVQIQLQKIRKQYQMLTMKDIIRDGHPTLRQAAELPLTE---EKETLIA 60
Db 2168 INLVLSVRETRQLTRQEQEKKAD-----EASEAACPKEDGLAQQQTQLN 2216
Query 61 MREFLYNSQDEELAKRIGLRSGVGLAQPQINISKRMIAVLIPDDGSGSKSYDMLVNPKIV 120
Db 2217 IRSLLVNPPEGFTMLNVSQS-----SERPLVHIEGSTTVFHSLASRLSIP 2267

RESULT 14
US-09-261-907-2
i Sequence 2, Application US/09261907A
i Patent No. 6294364

GENERAL INFORMATION:
; APPLICANT: ELLIS, CATHERINE
; APPLICANT: LONSDALE, JOHN
; APPLICANT: BERGSHA, DERK J.
; APPLICANT: MCONEY, JEFFREY L.
; APPLICANT: DEITERA, MEGAN E.
; APPLICANT: CHAPMAN, CONRAD
; TITLE OF INVENTION: HUMAN FAS
; FILE REFERENCE: GP-70603
; CURRENT APPLICATION NUMBER: US/09/261,907A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 2511
; TYPE: PRF
; ORGANISM: HOMO SAPiens
; US-09-261-907-2

Query Match 7.6%; Score 84; DB 4; Length 2511;
Best Local Similarity 21.6%; Pred. No. 9.3;
Matches 27; Conservative 32; Mismatches 40; Indels 26; Gaps 3;
Query 5 IANLFFVQIQLQKIRKQYQMLTMKDIIRDGHPTLRQAAELPLTE---EKETLIA 60
Db 2170 IANLVLSVRETRQLTRQEQEKKAD-----EASEAACPKEDGLAQQQTQLN 2218
Query 61 MREFLYNSQDEELAKRIGLRSGVGLAQPQINISKRMIAVLIPDDGSGSKSYDMLVNPKIV 120
Db 2219 IRSLLVNPPEGFTMLNVSQS-----SERPLVHIEGSTTVFHSLASRLSIP 2267

Query 121 SHSVQ 125
Db 2268 TYGLQ 2272

RESULT 15
US-08-557-122A-38
i Sequence 38, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Maland
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CURRENT ADDRESS:
; ADDRESS: No. 5879664 No. 5879664 Disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-367-9123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

i MOLECULE TYPE: peptide
US-08-557-122A-38

Search completed: July 11, 2003, 09:42:02
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:40:50 ; Search time 51 Seconds
(without alignments)

Title: US-09-896-580a-1

Perfect score: 1112
Sequence: 1 MYEYLNLFVIQKQIKR.....KDHPHQHTDAVEWHQQHHH 213

Scoring table: BLOSUM62 Gapext 0.5

Searched: 445758 seqs, 11641973 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```

1: /cgm2_6/ptodata/1/pubpaas/US07 NEW PUB.PEP:*
2: /cgm2_6/ptodata/1/pubpaas/PCT NEW PUB.PEP:*
3: /cgm2_6/ptodata/1/pubpaas/US06 PUBCOMB.PEP:*
4: /cgm2_6/ptodata/1/pubpaas/US07 PUBCOMB.PEP:*
5: /cgm2_6/ptodata/1/pubpaas/US07 PUBCOMB.PEP:*
6: /cgm2_6/ptodata/1/pubpaas/US08 NEW PUB.PEP:*
7: /cgm2_6/ptodata/1/pubpaas/US08 PUBCOMB.PEP:*
8: /cgm2_6/ptodata/1/pubpaas/US09 NEW PUB.PEP:*
9: /cgm2_6/ptodata/1/pubpaas/US09 NEW PUB.PEP:*
10: /cgm2_6/ptodata/1/pubpaas/US09 PUBCOMB.PEP:*
11: /cgm2_6/ptodata/1/pubpaas/US10 NEW PUB.PEP:*
12: /cgm2_6/ptodata/1/pubpaas/US10 PUBCOMB.PEP:*
13: /cgm2_6/ptodata/1/pubpaas/US60 NEW PUB.PEP:*
14: /cgm2_6/ptodata/1/pubpaas/US60 PUBCOMB.PEP:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1		41.1	203	10	US-09-862-005-2	Sequence 2, Appli
2		386.5	34.8	105	US-09-862-005-4	Sequence 4, Appli
3		161	14.5	216	9 US-10-156-761-10570	Sequence 10570, A
4		160.5	14.4	193	10 US-09-738-626-6508	Sequence 6508, Ap
5		146.5	13.2	224	9 US-10-156-761-14878	Sequence 2, Appli
6		143	12.9	169	9 US-09-738-626-5271	Sequence 14678, A
7		140	12.6	243	9 US-10-149-056-6	Sequence 5271, Ap
8		140	12.6	243	10 US-09-066-099-2	Sequence 6, Appli
9		139	12.5	186	9 US-10-156-761-14655	Sequence 2, Appli
10		98.5	8.9	210	9 US-10-149-226-4	Sequence 4, Appli
11		91	8.2	162	9 US-09-921-144-16	Sequence 16, Appli
12		91	8.2	163	9 US-09-921-144-14	Sequence 14, Appli
13		91	8.2	184	9 US-09-921-144-20	Sequence 20, Appli
14		91	8.2	185	9 US-09-921-144-18	Sequence 18, Appli
15		87.5	7.9	604	9 US-09-949-233-30	Sequence 30, Appli
16		86	7.7	522	9 US-09-949-79A-11	Sequence 11, Appli
17		84.5	7.6	604	9 US-09-949-293-26	Sequence 26, Appli
18		84	7.6	2509	9 US-10-237-271-1	Sequence 1, Appli

ALIGNMENTS

```

20 81.5 7.3 544 9 US-09-949-293-29
21 81 7.3 836 9 US-10-210-295-5
22 80 7.2 163 9 US-10-149-256-2
23 80 7.2 389 12 US-10-062-234-300
24 79 7.1 588 9 US-10-156-761-14205
25 78.5 7.1 600 10 US-09-974-592-12
26 78.5 7.1 604 9 US-09-949-293-28
27 78.5 7.1 604 9 US-09-953-0678-8
28 78 7.0 935 10 US-09-160-758-15
29 78 7.0 4349 9 US-10-160-758-15
30 77 6.9 468 10 US-09-925-300-1661
31 76 6.8 299 9 US-09-738-564-484
32 76 6.8 543 9 US-10-225-567A-426
33 75.5 6.8 599 9 US-09-949-293-31
34 75.5 6.8 602 9 US-09-953-0674-7
35 75.5 6.8 823 9 US-10-177-293-240
36 75.5 6.8 4999 9 US-09-976-059-15
37 75 6.7 215 9 US-10-172-399-2
38 75 6.7 1344 9 US-09-738-626-888
39 75 6.7 1979 9 US-10-205-823-419
40 75 6.7 350 9 US-10-197-340-266
41 74.5 6.7 350 9 US-09-251-225-3
42 74.5 6.7 350 9 US-10-205-823-341
43 74.5 6.7 385 9 US-09-925-29-966
44 74.5 6.7 385 9 US-10-106-698-4598
45 74.5 6.7 385 9 US-10-106-698-4598

```

```

SEQUENCE 1
US-09-862-005-2
Sequence 2, Application US/09862005
Filing Date: 21-May-2001
Prior Application Number: US/0202058196A1
Filing Date: 21-May-2001
Attorney/Agent Information: <Unknown>
Applicant: Lonetto, Michael A.
Title of Invention: def1
Number of Sequences: 5
Correspondence Address:
    Addressee: Dachart Price & Rhoads
    Street: 4000 Bell Atlantic Tower, 1717 Arch Street
    City: Philadelphia
    State: PA
    County: US
    Zip: 19103
Computer Readable Form:
Medium Type: Diskette
Computer: IBM Compatible
Operating System: DOS
Software: FASTSEQ for Windows Version 2.0
Current Application Data:
Application Number: US/09/862,005
Filing Date: 21-May-2001
Classification: <Unknown>
Prior Application Number: 08/991,023
Application Number: 08/991,023
Attorney/Agent Information: <Unknown>
Name: Falk, Stephan T
Registration Number: 36,795
Reference/Doctet Number: GM50010
Telecommunication Information:
    Telephone: 215-994-2488
    Telefax: 215-994-2222
    Telex: <Unknown>
Information for Seq ID No: 2:
Sequence Characteristics:
    Length: 203 amino acids
    Type: amino acid
    Strandedness: single
    Topology: linear
Sequence Description: SEQ ID NO: 2:
```

US-09-862-005-2

Query Match 41.1% ; Score 457.5; DB 10; Length 203;
 Best Local Similarity 50.3%; Pred. No. 2.4e-35;
 Matches 94; Conservative 37; Mismatches 49; Indels 7; Gaps 2;

Qy 18 KIRKVOYMLTMKDIDRQHPTLRQDAELSLPITKEKETLAMREFLYNSODEIARY 77
 Db 6 RITKAHHIDMNDTIRENPXPLRTVAEVTPLSDEILGEKMKQFLKHSODEVMAKM 65
 Qy 78 GERSGVGLAAPQINISKRMIAVLPD---MLYNPKVTSVSYQEAVL 130
 Db 66 GURGGVGLAAPQDLISKRIAVLVPNIVVEGETPOATEALMYPKVTSVSYQDAALG 125

131 TEGGCLSVDDNVAGVHVHENKTIKARDIEENDOLRKGYPAVQHEIDHNGMFD 190
 Db 126 BEBGCLSVDDNVPGVVRHARVIVDVFDRGEKERIKLGINSIVVQHEIDHNGMFD 185

Qy 191 HIDKDHPI 197
 Db 186 RINEKDP 192

RESULT 2

US-09-862-005-4
 Sequence 4, Application US/09862005
 Patent No. US201020158796A1
 GENERAL INFORMATION:
 APPLICANT: Lonetto, Michael A.
 TITLE OF INVENTION: def1
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dechert Price & Rhoads
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
 CITY: Philadelphia
 STATE: PA
 COUNTRY: US
 ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/862,005
 FILING DATE: 21-May-2001
 CLASSIFICATION: <Unknown>
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/991,023
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Falk, Stephan T.
 REFERENCE/DOCKET NUMBER: CMS0010
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-994-2488
 TELEFAX: 215-994-2222
 TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 155 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-862-005-4

Query Match 34.8%; Score 386.5; DB 10; Length 155;
 Best Local Similarity 54.2%; Pred. No. 7.9e-29;
 Matches 78; Conservative 28; Mismatches 31; Indels 7; Gaps 2;

Qy 61 MREFLYNSODEIARYGLRSVGLAAPQINISKRMIAVLPD---DGSGKSYD--M 113
 | :| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :

1 MNQFLKHSQDPVMAEKMGRLGGVGLAPQDLISKRIIAVLPNIVEGETPOAYDLEAI 60

Qy 114 LYNPCKVTSVSYQEAVLPTGEGCLSYDNTVAGLVERHNKTTIAYKLEGNDIQQLRKGYPA 173
 Db 61 MYNPCKVTSVSYQDNLGEEGCLSYDNTVAGLVERHNKTTIAYKLEGNDIQQLRKGYPA 120

Qy 174 IVFOHEIDHNGMFDIDKDP 197
 Db 121 IVQHEIDHNGMFDIDKDP 144

RESULT 3
 Sequence 15570, Application US/10156761
 Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHITAKI
 APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 10570

LENGTH: 216

TYPE: PR

ORGANISM: Streptomyces avermitilis
 US-10-156-761-10570

Query Match 14.5%; Score 161; DB 9; Length 216;

Best Local Similarity 30.2%; Pred. No. 2e-07;
 Matches 57; Conservative 30; Mismatches 80; Indels 22; Gaps 8;

Qy 11 VIGLQKIRKVQML-TMKDIDRSHPTQRKABELPLTK--BEKETLILAREFLIV 66
 Db 22 VIGADECEBERANERGTSRPTVQNPFLHKEKD---VTDGAELLEIV-A-DPA 75

Qy 67 NSQDEELAKRYGLRSVGLAAPQINISKRMIAVLPDGSKSYDYMVNPKVSYSEVOE 126
 Db 76 SQRAE-----GVGANQIGVLKVVDQDD-BGTHVGVCNPLVLPADR 125

Qy 127 AYL-PTEGCLSYDNTVAGLVERHNKTTIAYKLEGNDIQQLRKGYPA 185
 Db 126 RLDDSNEGCLSVPTAYPLA-RPDVAEVTQDERRGNPKVRETGPARCHQHETDHLYG 184

Qy 186 VMFYDHDK 194

Db 185 YLTDRASK 193

RESULT 4

US-09-739-626-6508

Sequence 6508, Application US/09738626
 Publication No. US20020197605A1

GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIJI
 APPLICANT: HAYASHI, MIKI
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TAMEISHI, NAOKO
 APPLICANT: SERNOH, AKIHIRO
 APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 PRIOR APPLICATION NUMBER: JP 99/377484
 CURRENT FILING DATE: 2000-12-18
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 6508
 LENGTH: 193
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-6508

Query Match 14.4%; Score 160.5; DB 9; Length 193;
 Best Local Similarity 29.7%; Pred. No. 1.9e-07; Mismatches 63; Indels 31; Gaps 9;
 Matches 52; Conservative 29; General Information: US/09/738,626
 PRIOR APPLICATION NUMBER: JP 00/159162
 CURRENT FILING DATE: 2000-12-18
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 6508

Qy 26 LTMIDITDGHPTLROKAELIPITKEKETILLAMREFLYNSQDEEIAKRY--GLRSG 82
 Db 1 MTVRPIVHGDPVLANPT---OLVTEP-----VSELQELIADMETMDVANG 44

Qy 83 VGLAAPQINISKRMIATIPDGSKSXTMVLNPKIVSSVQAYLTG----EGCIS 137
 Db 45 VGLAAQGVSPRIFTYCDPDD-EGVMERKGCFINPVLETSEIPEP-MPADGSDEEGCIS 102

Qy 138 V--DDNVAGLVERHNKTTIKADLEGNDIQLRLKGYPATFOEIDLHNGVNFYD 190
 Db 103 VPGSGFPPGRAH-WAKVT--GLNEKGENVSVBAGFLARFQHEVGHLDGFLYTD 154

RESULT 5
 US-09-825-345-2
 Sequence 2, Application US/09825345
 Patent No. US20020106750A1
 GENERAL INFORMATION:
 APPLICANT: FARWICK, MIKE
 APPLICANT: HUTHMACHER, KLAUS
 APPLICANT: BREHME, JENNIFER
 APPLICANT: PFEFFELBE, WALTER
 TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES WHICH CODE FOR THE def GENE
 FILE REFERENCE: 21123/280107/MAS
 CURRENT APPLICATION NUMBER: US/09/825,345
 PRIOR APPLICATION NUMBER: DE 100 46 228 6
 PRIOR FILING DATE: 2000-09-19
 PRIOR APPLICATION NUMBER: DE 101 13 957 8
 CURRENT FILING DATE: 2001-03-22
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 2
 LENGTH: 193
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-825-345-2

Query Match 14.4%; Score 160.5; DB 10; Length 193;
 Best Local Similarity 29.7%; Pred. No. 1.9e-07; Mismatches 63; Indels 31; Gaps 9;
 Matches 52; Conservative 29; General Information: US/09/738,626
 PRIOR APPLICATION NUMBER: JP 00/159162
 CURRENT FILING DATE: 2000-12-18
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 6508

Qy 26 LTMIDITDGHPTLROKAELIPITKEKETILLAMREFLYNSQDEEIAKRY--GLRSG 82
 Db 1 MTVRPIVHGDPVLANPT---OLVTEP-----VSELQELIADMETMDVANG 44

Qy 83 VGLAAPQINISKRMIATIPDGSKSXTMVLNPKIVSSVQAYLTG----EGCIS 137
 Db 45 VGLAAQGVSPRIFTYCDPDD-EGVMERKGCFINPVLETSEIPEP-MPADGSDEEGCIS 102

RESULT 6
 US-10-156-761-14878
 Sequence 14.8%; File Reference: US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAOJI
 APPLICANT: SAKAKI, YOSHIOKI
 APPLICANT: HATTORI, MASAHISA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-162
 CURRENT APPLICATION NUMBER: US/09/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-224089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO: 14878
 LENGTH: 224
 TYPE: PRT
 ORGANISM: Streptomyces avermitilis
 US-10-156-761-14878

Query Match 13.2%; Score 146.5; DB 9; Length 224;
 Best Local Similarity 30.0%; Pred. No. 4.7e-06;
 Mismatches 26; Indels 39; Gaps 9;
 Matches 54; Conservative 26; General Information: US/09/738,626
 PRIOR APPLICATION NUMBER: JP 2001-224089
 PRIOR FILING DATE: 2001-05-30
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO: 14878

Qy 31 LIRDGHTPLRKAAELIPITKEKETILLAMREFLYNSQDEEIAKRYGLRSGYGLAAPOQI 90
 Db 35 IVAADPDYTRGRGE---PYPGOLGPGLIA-RFV---EARLUTMHAAPFGTGLAAPOQV 83

Qy 91 NISKRMIAV----IPDD----GSGKSYDYMVNPKIVSSVQAYLTG----E 133
 Db 84 GVGLR-IAVTEPDAPYDPEEVAVRGRVPPFRVLPNP----SYAVGSDRAFFE 133

Qy 134 GCLSYDNYAVGILVHRNKTTIKADLEGNDIQLRLKGYPATFOEIDLHNGVNFYDHD 193
 Db 134 GCLSYDNYAVGILVHRNKTTIKADLEGNDIQLRLKGYPATFOEIDLHNGVNFYDHD 193

RESULT 7
 US-09-738-626-5271
 Sequence 5271, Application US/09738626
 Publication No. US2002017605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 5271
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5271

Query Match Score 143; DB 9; Length 169;
Best Local Similarity 24.6%; Pred. No. 6.9e-06;
Matches 43; Conservative 35; Mismatches 65; Indels 32; Gaps 6;

Qy 26 LTMKDIRCHPILRKQKAAELPLTKEKEETLAMREFLVNSQDEBIAKRYGIRSGVGL 85
Db 1 MAVREVRLFEDPVLADEV----VDFDESLSLTDMPDTMED-----AGSYGL 47

Qy 86 AAQONIISRMIAVLPIDGSKSIVDYMVNPKVLSASVQAYLP-----TG-EGLSLV 138
Db 48 AANQGVLRPFVY-----FDTSOEGERGRHGVINPWPFLTDQTKEKGCLSI 96

Qy 139 DDNVAGLVRHRNKLITIKAKDIEGNDIQLRLKGYPATVQEHEIDHNGMVFHDID 193
Db 97 ED-VSAETTRYETVRLSSQDRDGNPVGTVANGLLARCTQHETDHGVLFLRDL 150

RESULT 8 US-10-149-256-6

; Sequence 6 Application US/10149256
; Publication No. US2003000348811
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Shyam
; TITLE OF INVENTION: Regulated expression of Human Mitochondrial
; TITLE OF INVENTION: Deformylase
; FILE REFERENCE: 004974_00739
; CURRENT APPLICATION NUMBER: US/10/149, 256
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/EP00/12110
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US/10/149, 256
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: US/10/169, 615
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match Score 140; DB 9; Length 243;
Best Local Similarity 27.1%; Pred. No. 2.2e-05;
Matches 49; Conservative 32; Mismatches 55; Indels 44; Gaps 9;

Qy 35 GHPTIRQKAAELPLD-----PLTKEKEETLAMREFLVNSQDEBIAKRYGIRSGVGL 89
Db 71 GDPVLRGVAPVERAQLGPQLRQLTQRLY-----QWNER---RCYGLSAQ 115

Qy 90 INISKRMIAVLPID-----DGSGKSYDYM-----LYNPKVSHSVQAYLP-GFGC 135
Db 116 LGVPRQVLAELPALLCPRQRALRQMEPLRVRVNPSL---RVLDISRLVTPFGC 172

Qy 136 LSVDNTVAGL--VHRHKITIKAKDIEGNDIQLRLKGYPATVQEHEIDHNGMVFHDID 192
Db 173 ---ESTAGFLACYPFRQAVOISGLDENGBQVQASWAAIHQMDHLQGCLFIDRM 228

Qy 193 D 193
Db 229 D 229

RESULT 10 US-10-156-761-14555

; Sequence 14655, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIIKU
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: Novel Polynucleotides
; FILE REFERENCE: 249-362
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 14655
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis

Query Match Score 139; DB 9; Length 186;
Best Local Similarity 31.2%; Pred. No. 1.9e-05;

Qy 193 D 193
Db 229 D 229

RESULT 9

Matches 35; Conservative 24; Mismatches 49; Indels 4; Gaps 3;

Query 81 SGVGLAAPQINISRMIAVLIDDDSGKSYDMLVNPKVHSVSEAVLPTGGTCAVDD 140
 Best Local Similarity 8.2%; Score 91; DB 9; Length 162;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;
 Pred. No. 0.51;

Db 53 NGVLAANQIGVLRVYDCPDDSDVRVGH-VVNRLI--EADGVWLRGPGCULSLPG 109

Query 141 NVAGLVRHINKITKAKDIEGNDIQRLKGPAVIFQHEIDLNGVMFYDH 192
 Best Local Similarity 25.1%; Score 91; DB 9; Length 162;
 Matches 35; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 110 LEAG-TERIDRAVEGFTTDEPVRVLGFWPARCLOFECHDLDGIVYDRV 160

RESULT 11
 US-10-149-144-4
 ; Sequence 4, Application US/1049256
 ; GENERAL INFORMATION:
 ; APPLICANT: Ramakrishnan, Shyam
 ; TITLE OF INVENTION: Regulation of Human Mitochondrial
 ; TITLE OF INVENTION: Deformylase
 ; FILE REFERENCE: 004474-00739
 ; CURRENT APPLICATION NUMBER: US/10/149,256
 ; CURRENT FILING DATE: 2002-06-10
 ; PRIORITY FILING DATE: 2000-12-01 PCT/EP00/12110
 ; PRIOR APPLICATION NUMBER: US 60/239,106
 ; PRIOR APPLICATION NUMBER: US 60/169,615
 ; PRIOR APPLICATION NUMBER: 2000-10-11
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 4
 ; LENGTH: 210
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(210)
 ; OTHER INFORMATION: Xaa = Any Amino Acid

Query Match 8.2%; Score 91; DB 9; Length 210;
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 36 EKKEKEETLAMREFLYNSQDEIAK-RYGLRSVGAAPQNSRMLIPDDGSCK 108

Query 133 EGCLSVVDNAGLVRHINKITKAKDIEGNDIQRLKGPAVIFQHEIDLNGVMFYDH 192
 Best Local Similarity 25.1%; Score 91; DB 9; Length 162;
 Matches 36; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 3 3 EGCGESVAAFXA-CVPRFOXVQLISGELPNQEQVWQASGWAARIQHEMDHQCLFDKM 61

Query 193 D 193
 Best Local Similarity 25.1%; Score 91; DB 9; Length 162;
 Matches 36; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 62 D 62

RESULT 12
 US-09-921-144-16
 ; Sequence 16, Application US/09921144
 ; GENERAL INFORMATION:
 ; APPLICANT: ZHONG, Pingyu
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING CHIMERIC HETEROMULTIMERS
 ; FILE REFERENCE: 13403 0004.NEU00
 ; CURRENT APPLICATION NUMBER: US/09/921,144
 ; CURRENT FILING DATE: 2001-08-01
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 16
 ; LENGTH: 162
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Bluescript vector

Query Match 8.2%; Score 91; DB 9; Length 213;
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 83 -EAKLEKEERVSLRHLQSYGGCAQPHSKSKEVLMKCLL----- 83

Query 109 SYDMVNPKVSHSVQAEYLPTGBCLSYDNYAGLVHR-HINKITKAKDIEGNDIQR 167
 Best Local Similarity 25.1%; Score 91; DB 9; Length 162;
 Matches 35; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 84 -FAIPLVVP-FYSHSATS--RGGGTSRSLBGLQSENHLRKMTFLDXLEEVTMQL 138

Query 168 LKGYPALVPEHEI-DHNGVMFYDHDKHPLQHTDAVEYQHHH 213
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 138 DVGGCAAAYPYDVEDYAGG----- 163

RESULT 13
 US-09-921-144-14
 ; Sequence 14, Application US/09921144
 ; Publication No. US20030027247A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZHONG, Pingyu
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING CHIMERIC HETEROMULTIMERS
 ; FILE REFERENCE: 13403 0004.NEU00
 ; CURRENT APPLICATION NUMBER: US/09/921,144
 ; CURRENT FILING DATE: 2001-08-01
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 14
 ; LENGTH: 163
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Bluescript vector

Query Match 8.2%; Score 91; DB 9; Length 163;
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 36 EKKEKEETLAMREFLYNSQDEIAK-RYGLRSVGAAPQNSRMLIPDDGSCK 108

Query 109 SYDMVNPKVSHSVQAEYLPTGBCLSYDNYAGLVHR-HINKITKAKDIEGNDIQR 167
 Best Local Similarity 25.1%; Score 91; DB 9; Length 162;
 Matches 35; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 84 -FAIPLVVP-FYSHSATS--RGGGTSRSLBGLQSENHLRKMTFLDXLEEVTMQL 138

Query 168 LKGYPALVPEHEI-DHNGVMFYDHDKHPLQHTDAVEYQHHH 213
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 139 DVGGCAAAYPYDVEDYAGG----- 163

RESULT 14
 US-09-921-144-20
 ; Sequence 20, Application US/09921144
 ; Publication No. US20030027247A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZHONG, Pingyu
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING CHIMERIC HETEROMULTIMERS
 ; FILE REFERENCE: 13403 0004.NEU00
 ; CURRENT APPLICATION NUMBER: US/09/921,144
 ; CURRENT FILING DATE: 2001-08-01
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 20
 ; LENGTH: 184
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence

Query Match 8.2%; Score 91; DB 9; Length 163;
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 36 EKKEKEETLAMREFLYNSQDEIAK-RYGLRSVGAAPQNSRMLIPDDGSCK 108

Query 109 SYDMVNPKVSHSVQAEYLPTGBCLSYDNYAGLVHR-HINKITKAKDIEGNDIQR 167
 Best Local Similarity 25.1%; Score 91; DB 9; Length 162;
 Matches 35; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 84 -FAIPLVVP-FYSHSATS--RGGGTSRSLBGLQSENHLRKMTFLDXLEEVTMQL 138

Query 168 LKGYPALVPEHEI-DHNGVMFYDHDKHPLQHTDAVEYQHHH 213
 Best Local Similarity 25.1%; Pred. No. 0.51;
 Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;

Db 139 DVGGCAAAYPYDVEDYAGG----- 163

FEATURE:
; OTHER INFORMATION: Bluescript vector
US-09-921-144-20

Query Match 8.2%; Score 91; DB 9; Length 184;
Best Local Similarity 25.1%; Pred. No. 0.6;
Matches 42; Conservative 27; Mismatches 56; Indels. 42; Gaps 8;

Qy 50 LTKKEKETLAMREVLQNSQDEETAK-RYGRSGVGLAQPQINISKRMIAVLIPDDGSK 108
Db 35 LEKENRL---EKIAKEEERSEERHQLOQS/VGCCGPQFHSKEVIMKLL----- 82

Qy 109 SYDMLNPKIVSHSVOEAYLPTEGGCLSYVDNYAGLVR-HNKLTAKADIEGNDIQR 167
Db 83 -FAIPLVVP-FYSSATRS---RGGGTSRLEQGSENRNLNKTFSLDDLEEVTMOLQ 137

Qy 168 LKGYPALVEQHEI-DHNGYMFHDIDKDHPLQPTDAVEHQQHHH 213
Db 138 DVGGCAAAVPDVPPYAGG----- 162

RESULT 15

US-09-921-144-18
; Sequence 18, Application US/0921144
; Publication No. US20030027247A1

; GENERAL INFORMATION:
; APPLICANT: ZHONG, Pingyu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING
; CHIMERIC HETEROTIMERS
; FILE REFERENCE: 13403_004.NPUS00
; CURRENT APPLICATION NUMBER: US/09/921,144
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 18
; LENGTH: 185
; TYPE: PRT
; FEATURE:
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Bluescript vector

US-09-921-144-18

Query Match 8.2%; Score 91; DB 9; Length 185;
Best Local Similarity 25.1%; Pred. No. 0.61;
Matches 42; Conservative 27; Mismatches 56; Indels. 42; Gaps 8;

Qy 50 LTKKEKETLAMREVLQNSQDEETAK-RYGRSGVGLAQPQINISKRMIAVLIPDDGSK 108
Db 36 LEKENRL---EKIAKEEERSEERHQLOQS/VGCCGPQFHSKEVIMKLL----- 83

Qy 109 SYDMLNPKIVSHSVOEAYLPTEGGCLSYVDNYAGLVR-HNKLTAKADIEGNDIQR 167
Db 84 -FAIPLVVP-FYSSATRS---RGGGTSRLEQGSENRNLNKTFSLDDLEEVTMOLQ 138

Qy 168 LKGYPALVEQHEI-DHNGYMFHDIDKDHPLQPTDAVEHQQHHH 213
Db 139 DVGGCAAAVPDVPPYAGG----- 163

Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: July 11, 2003, 03:33:30 ; Search time 40 Seconds
 (without alignments)
 511.916 Million cell updates/sec

Title: US-09-896-580A-1
 Perfect score: 1112

Sequence: 1 MYEMLNLFTVQLQKIR.....KDHPHQPHTDAAEVHQHHH 213

Scoring table: BLOSUM62

Gapop 10.0 , Gappext 0.5

Searched:

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries

Database : PIR 73;*

1: Pirl1;*

2: Pirl2;*

3: Pirl3;*

4: Pirl4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	84.4	183	2 A89879	hypothetical prote
2	575.5	51.8	184	1 D69862	formylmethionine d
3	536.5	48.2	183	2 AB1563	formylmethionine d
4	533.5	48.0	183	2 AC1206	formylmethionine d
5	506.5	45.5	182	2 383982	formylmethionine d
6	462.5	41.6	203	2 E98035	formylmethionine d
7	461.5	41.5	203	2 E95169	polypeptide deform
8	435	39.2	211	2 H86694	polypeptide deform
9	273.5	24.6	198	2 P82896	hypothetical prote
10	263	23.7	198	2 A90528	hypothetical prote
11	230	20.7	216	2 S73913	polypeptide deform
12	226.5	20.4	224	2 H97709	formylmethionine d
13	222.5	20.0	226	1 G64211	formylmethionine d
14	217.5	19.6	168	2 C82394	polypeptide deform
15	215.5	19.4	191	2 AI2056	polypeptide deform
16	211.5	19.0	169	1 D64082	ins protein homolo
17	207	18.6	181	2 C81680	N-formylmethionyla
18	206	18.5	273	2 T48639	polypeptide deform
19	205.5	18.5	232	2 H75274	hypothetical prote
20	203	18.3	181	2 B71536	probable polypepti
21	198	17.8	164	2 C72224	polypeptide deform
22	193.5	17.4	170	2 F97403	polypeptide deform
23	193.5	17.4	170	2 AF2621	formylmethionine d
24	190	17.1	173	2 E84987	N-formylmethionyl-
25	189	17.0	150	2 A97205	polypeptide deform
26	184.5	16.6	160	1 F9613	formylmethionine d
27	184.5	16.6	169	2 AB1010	hypothetical prote
28	184	16.5	259	2 E86288	

Summary

%

Query Match

Length

ID

Description

RESULT 1	A89879	hypothetical protein pdf1 [Imported] - <i>Staphylococcus aureus</i> (strain N315)
C;Species:	Staphylococcus aureus	
C;Date:	10-May-2001 #sequence_revision 10-May-2001 #text_change 31-Dec-2001	
C;Accession:	A89879	
R;Kuroda, N.; Ohta, T.; Uchiyama, I.; Baba, T.; Kobayashi, I.; Cui, L.; Ogi		
ma, A.; Miutani-U, Y.; Inoue, R.; Savano, T.; Kaito, C.; Sekimizu, K.		
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.		
Lacet, 357, 1225-1240, 2001		
A;Title: Whole Genome sequencing of methicillin-resistant <i>Staphylococcus aureus</i> .		
A;Reference number: A89879; PMID:21311952; PMID:11418146		
A;Accession: A89879		
A;Status: preliminary		
A;Molecule type: DNA		
A;Residues: 1-183 <KUR>		
A;Cross-references: GB:BA000018; PID:913700892; PID:BA842188-1; GSPDB:GN00149		
A;Experimental source: strain N315		
C;Genetics:		
A;Gene: pdf1		
C;Superfamily: polypeptide deformylase		

RESULT 1

Best Local Similarity 84.4%; Score 938; DB 2; Length 183;

Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 25 MLTMKDIIRDGHPTLROKAAPALEPLTKEKEETLIAAMREFLYNSDDEEIAKRYGLRSGVG 84

Db 1 MLTMKDIIRDGHPTLROKAAPALEPLTKEKEETLIAAMREFLYNSDDEEIAKRYGLRSGVG 60

Qy 85 LAAPQINITSKRMIAVILIPDGGSGSYDYLMLVNPKVISHSYQEAYLPTGEGCLSYDDNVAG 144

Db 61 LAAPQINITSKRMIAVILIPDGGSGSYDYLMLVNPKVISHSYQEAYLPTGEGCLSYDDNVAG 120

Qy 145 LVRENKNTIKAKIDENDIQRLKGKPAIVFHEIDHLNGVMFDHIXDKHLOPHDIA 204

Db 121 LVRENHNTIKAKIDENDIQRLKGKPAIVFHEIDHLNGVMFDHIDKHNFLQPHDIA 180

Qy 205 VEV 207

Db 181 VEV 183

RESULT 2

D69862

formylmethionine deformylase homolog ykrB - *Bacillus subtilis*

C;Species: *Bacillus subtilis*

C;Accession: D69862

R;Kunst, F.; Ogasawara, N.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Cr

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallez, J.; Harwood, C.R.; Henaut, A.; Hillert, H.; Holzapfel, S.; Hosono, S.; Hull, M.F.; Koettler, P.; Konigstein, G.; Krogh, S.; Kumada, M.; Kurita, K.; Lapidus, A.; Lardinois, S.M.; Levine, A.; Liu, H.J.; Masuda, S.; Matay, Y.; Ogawa, K.; Ogiwara, A.; Outcalt, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scuffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, R.; Winters, P.; Widat, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Totaro, V.; Uchiyama, K.; Authors: Yoshikawa, H.P.; Zumstein, B.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A;Reference number: Ab6580; MUID:28044033; PMID:9384377
 A;Accession: D69862
 A;Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-184 <GLA>
 A;Cross-references: GB:Z99111; GB:AL009126; NID:92633699; PIDN:CAB13329.1; PID:92633827
 A;Experimental source: strain 168
 C;Genetics:
 C;Superfamily: polypeptide deformylase
 Query Match 51.8%; Score 505.5; DB 1; Length 184;
 Best Local Similarity 59.5%; Pred. No. 3e-39;
 Matches 110; Conservative 33; Indels 1; Gaps 1;
 C;Species: *Listeria innocua*
 C;Accession: AB1563
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Bagner, F.; Berche, P.; Bloeckel, R.; Dominguez-Bernal, G.; Buchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihi, F.D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitouram, A.; Ok, C.; Schluerer, T.; Simoes, N.; Tierre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C;Title: Comparative genomics of *Listeria* species
 A;Accession: AC1206
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-183 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CAC99129.1; PID:916410453; GSDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 C;Superfamily: polypeptide deformylase
 Query Match 48.0%; Score 533.5; DB 2; Length 183;
 Best Local Similarity 57.5%; Pred. No. 7.2e-36;
 Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;
 C;Species: *Listeria innocua*
 C;Accession: AB1563
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Bagner, F.; Berche, P.; Bloeckel, R.; Dominguez-Bernal, G.; Buchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihi, F.D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitouram, A.; Ok, C.; Schluerer, T.; Simoes, N.; Tierre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C;Title: Comparative genomics of *Listeria* species
 A;Accession number: AB1077; MUID:2153279; PMID:11679669
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-183 <GLA>
 A;Cross-references: GB:Z99111; GB:AL009126; NID:92633699; PIDN:CAB13329.1; PID:92633827
 A;Experimental source: strain 168
 C;Genetics:
 C;Superfamily: polypeptide deformylase
 Query Match 205 VEHQ 209
 Db 180 TAER 184
 C;Species: *Listeria innocua*
 C;Accession: AB1563
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Bagner, F.; Berche, P.; Bloeckel, R.; Dominguez-Bernal, G.; Buchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihi, F.D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitouram, A.; Ok, C.; Schluerer, T.; Simoes, N.; Tierre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C;Title: Comparative genomics of *Listeria* species
 A;Accession number: AB1077; MUID:2153279; PMID:11679669
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-183 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CRG96274.1; PID:916413502; GSDB:GN00178
 A;Experimental source: strain Clip11262
 C;Genetics:
 C;Superfamily: polypeptide deformylase
 Query Match 48.2%; Score 536.5; DB 2; Length 183;
 Best Local Similarity 58.6%; Pred. No. 4.1e-36;
 Matches 106; Conservative 28; Mismatches 44; Indels 3; Gaps 2;

RESULT 4
 AC1206
 C;Species: *Listeria monocytogenes*
 C;Sequence: Listeria monocytogenes deformylase and to *B. subtilis* YkrB protein homolog lmo1051 [imported]
 C;Accession: AC1206
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Bagner, F.; Berche, P.; Bloeckel, R.; Dominguez-Bernal, G.; Buchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihi, F.D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitouram, A.; Ok, C.; Schluerer, T.; Simoes, N.; Tierre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C;Title: Comparative genomics of *Listeria* species
 A;Accession number: AB1077; MUID:2153279; PMID:11679669
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-183 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CAC99129.1; PID:916410453; GSDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 C;Superfamily: polypeptide deformylase
 Query Match 48.0%; Score 533.5; DB 2; Length 183;
 Best Local Similarity 57.5%; Pred. No. 7.2e-36;
 Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;
 C;Species: *Listeria innocua*
 C;Accession: AB1563
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Bagner, F.; Berche, P.; Bloeckel, R.; Dominguez-Bernal, G.; Buchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihi, F.D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitouram, A.; Ok, C.; Schluerer, T.; Simoes, N.; Tierre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, C;Title: Comparative genomics of *Listeria* species
 A;Accession number: AB1077; MUID:2153279; PMID:11679669
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-183 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CRG96274.1; PID:916413502; GSDB:GN00178
 A;Experimental source: strain Clip11262
 C;Genetics:
 C;Superfamily: polypeptide deformylase
 Query Match 48.2%; Score 536.5; DB 2; Length 183;
 Best Local Similarity 58.6%; Pred. No. 4.1e-36;
 Matches 106; Conservative 28; Mismatches 44; Indels 3; Gaps 2;

RESULT 5
 B83982
 Formylmethionine deformylase B82658 [imported] - *Bacillus halodurans* (strain C-125)
 C;Species: *Bacillus halodurans*
 C;Accession: B83982
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 R;Itakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Mastui, N.; Fuji, F.; His
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* an
 A;Accession number: A83550; MUID:20512582; PMID:11058132
 A;Accession: B83982
 A;Status: preliminary

25 MLTMIDIRDGHTIYRQKAAELEPLTKEETLILAREVLYNSQDEETAKRYGLRSQVG 84

A;Molecule type: DNA			
A;Residues: 1-182 <STO>			
A;Cross-references: GB:AP001516; GB:BA000004; NID:910175192; PIDN:BAB06377.1; GSPDB:GN001			
A;Experimental source: strain C-125			
C;Genetics:			
A;Gene: BH2658	C;Superfamily: polypeptide deformylase		
Query Match 45.5% Score 506.5; DB 2; Length 182;			
Best Local Similarity 58.0%; Pred. No. 1.1e-33;			
Matches 98; Conservative 31; Mismatches 39; Indels 1; Gaps 1;			
Qy 25 MIFTMDITRGCHPTLROKAALEELDTKEKEETLLARNEFLVNSDENTARYGRSGVG 84			
Db 1 MIFTMDITRGCHPTLROKAALEELDTKEKEETLLARNEFLVNSDENTARYGRSGVG 60			
Qy 85 LAAPQINISRMIAVLIPDDEGGSKYDYMILNKPLISHVQEAALPLTGGCLSVDDNVAG 144			
Db 61 LAAPQIGLSQMVAHTTDENE-KEYSVLIVENKLTIESSTMTHEGEGGLSVDREVGQ 119			
Qy 145 IYHRHAKITKAIDKEDNDIQLRKLCPAIVFQEIDHLNEMFYDHD 193			
Db 120 IYPRHAKITKAIDKEDNDIQLRKLCPAIVFQEIDHLNEMFYDHD 168			
RESULT 6			
B98035 C-terminal methionine deformylase (EC 3.5.1.31) fms [imported] - Streptococcus pneumoniae (strain TIGR4)			
C;Species: Streptococcus pneumoniae			
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001			
C;Accession: E98035			
A;Bioskins, J.A.; Alborn Jr., W.; Arnold, J.; Biaszczak, L.; Burgett, S.; DeHoff, B.S.; Ee, R.; LeBlanc, D.J.; Lee, I.N.; Larkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M.Y.; Sun, P.M.; Winkler, M.B.			
A;J. Bacteriol. 183: 5709-5717, 2001			
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.			
A;Reference number: A97872; MUID:21429245; PMID:11544234			
A;Accession: E98035			
A;Residues: 1-203 <KTR>			
A;Cross-references: GB:AE007317; PIDN:AA00114.1; PID:915458954; GSPDB:GN00174			
C;Genetics:			
A;Gene: fms	C;Keywords: hydrolase		
Query Match 41.6% Score 462.5; DB 2; Length 203;			
Best Local Similarity 50.8%; Pred. No. 4.3e-30;			
Matches 95; Conservative 37; Mismatches 48; Indels 7; Gaps 2;			
Qy 18 KIRKVQYMLMKDILRDGHPTLROKAALEELDTKEKEETLLARNEFLVNSDEEALKRY 77			
Db 6 RETKAHLIMNDIRECNPRLRTVAEFTPLSDQETILGKAKAQFLHSQPWAEKM 65			
Qy 78 GRSGYCLAAPOQINISKMTAVIIPD----DSGSKSYD-Y-MLYNPXKTVSEVQDYLPL 130			
Db 66 GLRGYCLAAPOQINISKMTAVIIPD----DSGSKSYD-Y-MLYNPXKTVSEVQDYLPL 125			
Qy 131 TGEGLSVDDNVAGLVEHNAKATIKADIEGNDIQLRKGPAIVFQEIDHLNEMFYD 190			
Db 126 EGEGGLSVDDNVPGYVVRHARVTVDYXFDKDGEXRIKUGTGINIVQHEIDHLNEMFYD 185			
Qy 191 HIDKDHP 197			
Db 186 RINEKD 192			
RESULT 7			
E95169 C-terminal methionine deformylase [imported] - Streptococcus pneumoniae (strain TIGR4)			
C;Species: Streptococcus pneumoniae			
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001			

C;Accession: E95169	R;Teitelbaum, J.D.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heijne, H.; Umayam, L.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.;on, T.; Hickney, B.K.; Holt, I.E.	A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrissey, A.;Title: Complete Genome Sequence of a virulent isolate of <i>Streptococcus pneumoniae</i> . A;Reference number: A95000; MUID:21357209; PMID:11463916	A;Accession: E95169	A;Status: preliminary
A;Residues: 1-203 <STO>	A;Cross-references: GB:ME05672; PIDN:AAK7550.1; PID:914972945; GSPDB:GN00164; TIGR:ST: C;Genetics:	A;Experimental source: strain TIGR44	A;Gene: SP1456	
Query Match 41.5%; Score 461.5; DB 2; Length 203;	Best Local Similarity 50.8%; Pred. No. 5.2e-30;	Matches 95; Conservative 37; Mismatches 48; Indels 7; Gaps 2;		
Qy 18 KIRKVQMLTKDITRDGHPTRQXAEELPLTKEERETLJAMREFLYNSQDEEBILARY 77	db 6 RITKAHLIMNDLIRENPNTURATAEVTPLSDQETLGEKMRQFLKESQDPVAMEK 65			
Qy 78 GLRSGYGLAQPQINISKRMIAVLPID---DGSKGSKSYD -MLYNPKTVYSHSYQEAQYLP 130	Db 66 GARGGGYGLAQPQDLISKRIALVLPNIVEEGEGTQEAYDLEAIMTPKTVYSHSYQDAALG 125			
Qy 131 TEGGLSVSYDNVAGLVERHNKTITRKDLEGNDIQLRLKGPAFAIYQHEIDHNGMVFYD 190	Db 126 EGEGGLSVYDNEPEGVYVHRVTDYFDKDGEXERIKLKGNTIVYQHEDHNGMVFYD 185			
Qy 191 HDKDHP 197	Db 186 RINEKDP 192			
RESULT: 8				
H86694	polypeptide deformylase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)			
C;Species: Lactococcus lactis subsp. lactis	C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001			
C;Accession: H86694	R;Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, P.;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, P.; Genome Res. 11, 731-753, 2001			
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. ssp. Res. 11, 731-753, 2001	A;Reference number: A86625; MUID:21335186; PMID:11337471			
A;Accession: H86694	A;Molecule type: DNA			
A;Status: preliminary	A;Residues: 1-211 <STO>			
A;Cross-references: GB:AB005176; PIDN:AAK04658.1; PID:912723447; GSPDB:GN00146	A;Experimental source: strain IL1403			
C;Genetics:	A;Gene: def			
Query Match 39.2%; Score 436; DB 2; Length 211;	Best Local Similarity 47.9%; Pred. No. 6.2e-28;	Matches 93; Conservative 38; Mismatches 51; Indels 12; Gaps 3;		
Qy 16 QIKTRQVQMLTKDITRDGHPTRQXAEELPLTKEERETLJAMREFLYNSQDEEBILARY 75	Db 7 QANLAKHISMDDITRGYPPLREAVNTPLSDEDITLGERMLQFLHNSODPVAE 66			
Qy 76 RYGRSEVGGLAQPAQINISKRMIAVLP----DDGS----GKSSTDY-MLVNPKTVSHS 123	Db 67 KMGLRGCVGAQNLGLKKVAVLPNVEBDGNEIPPEKAYCAREMYNAKVSHS 126			
Qy 124 VOEAYLPTGEGCLSYDNNVAGLVERHNKTITRKDLEGNDIQLRLKGPAFAIYQHEIDH 183	Db 127 VQDAVEEGGLSYDREPGTVVRHARVTEYTKNEGEKKIRLQDFPAICVHEIDHT 186			

Query Match 20.4%; Score 226.5; DB 2; Length 224;
 Best Local Similarity 34.1%; Pred. No. 5e-11; Gaps 9;
 Matches 72; Conservative 29; Mismatches 73; Indels 37; Gaps 9;

Qy 12 IOLKOKIRKVQYMLTMKDIDRGHPTLROKAAELEPLTKBS-KETLIAAREFLVNSQD 70
 29 ISLPQTNTIQTNSSED-----KTRIKAKTINFPLSSEDLRISILEKKCY---DQE 79

Db 71 EBIARRYGRSGVGLAAQOINISKRMATAVLIPDDGSKSY--DYLNVNPKVHSVQDAY 128
 80 ENCA-----GLAAPQIGSKCITIFAYEABLKWKHPDLCMTPKTI-WINSY 128

Qy 129 LPTG-----EGCLSYDNDVAGLVRHKNTIKITAKDIEGNDIOLRILKGPAIVFOHQED 181
 Db 129 XPIGDIXKHEDYEGGSV-ENATGPVAFRKCHYTAHDINGNQIAGEGFLARYTQHEID 187

Qy 182 HNGNMFDYHTI-----DRDHPLQPHPTDAY 206
 Db 188 HNGKVFELDVAPRKIMKXEELEMRKKAME 218

RESULT 13
 G64211 for mylmetheonine deformylase homolog - Mycoplasma genitalium
 C;Species: Mycoplasma genitalium
 C;Date: 10-Sep-1999 #text_change 07-Dec-1999
 C;Accession: G64211
 R.;Frueh, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
 M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Sauder, D.M.; Phillips, C.A.; Merrick, J.
 'C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A;Title: The minimal gene complement of Mycoplasma genitalium.
 A;Reference number: G64211
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Cross-references: GB:J39690; PMID:7569393
 A;Experimental source: strain G-37
 C;Genetic code: SGC3
 A;Start codon: GTG
 C;Superfamily: polypeptide deformylase

Query Match 20.0%; Score 222.5; DB 1; Length 226;
 Best Local Similarity 31.5%; Pred. No. 1.e-10; Gaps 5;
 Matches 64; Conservative 35; Mismatches 95; Indels 9; Gaps 5;

Qy 1 MYEYLNLLFTVIQ--LQKIKIRKVQYMLTMKDIDRGHPTLROKAAELEPLTKERETL 58
 11 MTKILPWLFTSIVRLTTLFSLMTFOPPKWLYFDDNLINKPTEAVNPF-DEQIETC 69

Db 59 TAMREFIVNSSDEELAKRGLRSGVGLAAQOINISKRMATAVLIPDDGSKSYDNYMLNPK 118
 70 IKKMIAYDASYDGAQEDILPFIGIGANAOIGYWKQLFYHIND--INNEKKCCLINEK 127

Qy 119 IVSHSYCEAVLPTPGEGLCSVDNDVAGHVHRNKTTKAD-TEGNDIOLRILKGPAIVFQ 177

Db 128 IIDQSENKAFLESGGCLSVKCOHKGVTSWITKGTDWEKKBTITATGLFGMCLQ 187

Qy 178 HEDHNGVMFYDHDKEPLQ 200
 Db 188 HEDHNLGRFFYQRI--NPLNP 207

RESULT 14
 C82294 polypeptide deformylase VCA0150 [Imported] - Vibrio cholerae (strain N16961 serogroup O1
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_change 20-Aug-2000
 C;Accession: C82294
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vanathavan, J.J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20405833; PMID:10952301
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-168 <HEL>
 A;Cross-references: GB:AE04356; GB:AE04353; NID:99657536; PID:NAAF96063.1; GSPDB:GN
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VCA0150
 A;Map position: 2
 C;Superfamily: polypeptide deformylase

Query Match 19.6%; Score 218.5; DB 2; Length 168;
 Best Local Similarity 31.4%; Pred. No. 1.5e-10; Gaps 6;
 Matches 58; Conservative 39; Mismatches 65; Indels 23; Gaps 6;

Qy 26 IWTMDIDRGHPTLROKAAELEPLTKERETLAMREFLYNSQDEETAKRVLGRSGVGL 85
 Db 1 MAVIETLPPDERLRLYQSKY-----TDAVSQVTLI-----DDLDLTIVATDNGIGL 47

Qy 86 AAPOQINTSKRMATAVLIPDDGSKSYDNYMLNPKTVSHTOBAYLPTGSGSCLSYDDVNAGL 145
 Db 48 AAPOYG----REPAVIVYDLSNDRDQPLVLPINPKVGSNM---GQGSCSVPDXAD- 100

Qy 146 VHRENKITHKADIEGNDIOLRILKGPAIVFQEHIDHLNVMFYDHDIDKHDPLQPTDAV 205
 Db 101 VERITSVIVTEALDREGKPKRIETSDFLAVIOMEHDLSGMLFIDYLs--PLKQQMAMK 157

Qy 206 EVFHQ 210
 Db 158 KVXCH 162

RESULT 15
 A12056 polypeptide deformylase [Imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp.
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. Strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_change 30-Jun-2002
 C;Accession: A12056
 R;Kansko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoo, M.; Takarawa, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium PCC 7120
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Status: preliminary
 A;Molecule type: DNA
 A;Cross-references: GB:J39690; PMID:91045782; PID:91045785; TIGR:NG106
 C;Genetic code: SGC3
 A;Start codon: GTG
 C;Superfamily: polypeptide deformylase

Query Match 19.6%; Score 217.5; DB 2; Length 179;
 Best Local Similarity 33.1%; Pred. No. 2e-10; Gaps 5;
 Matches 57; Conservative 30; Mismatches 60; Indels 25; Gaps 5;

Qy 31 IIRDGHPTLROKAAELEPLTKERETLAMREFLYNSQDEETAKRVLGRSGVGLAQPQI 90
 Db 8 IIQLGNPPLRQAAW-----NHDATQQLDIIATVAKAVGVTASPOV 55

Qy 91 NIISERMLTAVLIPDDGSKSYDNYMLNPK 143
 Db 56 AQSTRLFVA--SRPNPRYPHADEMPPTAMNPKVGHSTE--IVEWEGGLSLV-PGIR 109

Qy 144 GLVERHNKTTIKAKDGTGNDIOLRILKGPAIVFQHEIDHLNVMFYDHDIDK 195

Db 110 GLVPRHQAIIEWYTDYGNLQKOTLTDFVARIFQHEFDHLDGVLFIDRVESN 161

Search completed: July 11, 2003, 09:41:27
Job time : 41 secs

GenCore version 5.1.6						
Copyright (c) 1993 - 2003 Compugen Ltd.						
MM protein - protein search, using sw model						
run on:	July 11, 2003, 09:26:10	search time	23 Seconds			
		(without alignment)				
		384,107 Million cell updates				
title:	US-09-896-580A-1					
perfect score:	11.12					
sequence:	I MYEYLNLLFTVIQLQIKIR.....					KDHPLOPHTDAVEVHQHHH 21
scoring table:	BLOSUM62					
searched:	Gapop 10.0 , Gapext 0.5					
total number of hits satisfying chosen parameters:	112892					
minimum DB seq length:	0					
maximum DB seq length:	200000000					
post-processing:	Minimum Match 0%					
	Maximum Match 100%					
	Listing First 45 summaries					
database :	Swissprot_40:					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.						
SUMMARIES						
%						
result No.	Score	Query	Match	Length	DB	ID
1	575.5	51.8	184	1	DEF2_BACSU	
2	540.5	48.6	184	1	DEF2_BACST	
3	506.5	45.5	182	1	DEF_BACHD	
4	463.5	41.5	203	1	DEF_SRPN	
5	426	38.3	196	1	DEF_IACLA	
6	230	20.7	216	1	DEF_MCPN	
7	227	20.4	170	1	DEF_PASMU	
8	222.5	20.0	216	1	DEF_MTGB	
9	215.5	19.4	192	1	DEF_TETM	
10	211.5	19.0	169	1	DEF_HAEIN	
11	208	18.7	273	1	DEF_C_ARATH	
12	207	18.6	181	1	DEF_CHLMU	
13	203	18.3	181	1	DEF_CHLTR	
14	198	17.8	164	1	DEF_IHEM	
15	190	17.2	173	1	DEF_BTCAI	
16	188	16.9	277	1	DEF_LYCES	
17	187	16.8	279	1	DEF_C_LYCES	
18	184.5	16.6	160	1	DEF1_BACSU	
19	184	16.5	259	1	DEF_PARATH	
20	182.5	16.4	169	1	DEF_AQAB	
21	180.5	16.2	168	1	DEF_ECOLI	
22	180.5	16.2	187	1	DEF_STNV3	
23	178.5	16.1	150	1	DEF_COLAB	
24	177.5	16.0	175	1	DEF_CANDE	
25	176.5	15.9	187	1	DEF_FREDDI	
26	176	15.8	170	1	DEF_XYLPA	
27	175	15.7	168	1	DEF_PSEAE	
28	175	15.7	186	1	DEF_CHEPN	
29	169.5	15.2	169	1	DEF_VILCH	
30	163	14.7	167	1	DEF_NEIMA	
31	155	13.9	175	1	DEF_RICON	
32	151.5	13.6	174	1	DEF_HALLC	
33	148.5	13.4	162	1	DEF_TREPA	

RESULT 1
 DEF2_BACSU
 ID DEF2_BACSU STANDARD; PRT; 184 AA.
 AC Q45195;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Peptid deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase 2).
 DE DEFPS.
 GN OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1423;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Caldwell R.M., Ferrari B.;
 RT "Sequence analysis of the mobA-amps region of the *Bacillus subtilis* chromosome.",
 RT Submitted (JUL-1997) to the EMBL/GenBank/DDBU databases.
 RN [2] SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogassawa N., Mosser L., Albertini A.M., Alloni G.,
 RA Azereedo V., Belterro M.G., Bessieres P., Bolotin A., Borchart S.,
 RA Borissi R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brionnet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Comerton J.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Ellmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferriger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S., Glaser P., Gofteau A., Golightly B.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Hoisappel S., Hosono S., Hullu M.F., Itaya M., Jones L.,
 RA Joris B., Karancka D., Kasahara Y., Klaerz-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauver J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Muel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Oiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portebele D., Porwollik S., Prescott A.M.,
 RA Presecan B., Pujic C., Purcell B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi T., Sekowska A., Seror S.J., Sevor P., Shin S.-B., Soldo B.,
 RA Soriano A., Tacconi E., Takagi T., Takashii H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandembol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedder E., Weitberg T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H., Zumstein B., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.",
 RT P57877 pasteurella st
 RN

ALIGNMENTS

CHARACTERIZATION, SEQUENCE FROM N.A., AND CHARACTERIZATION.

RP MEDLINE-21322705; PubMed=11429456;
 RX Haas M.; Beyer D.; Gahlmann R.; Freiberg C.;
 RA "YcbB is the main Peptide deformylase in *Bacillus subtilis*, a
 subacterium containing two functional Peptidyl deformylases.";
 RT Microbiology 147:1783-1791 (2001).

!-: FUNCTION: Removes the formyl group from the N-terminal Met of
 newly synthesized proteins. Requires at least a dipeptide for an
 efficient rate of reaction. N-terminal L-methionine is a
 prerequisite for activity but the enzyme has broad specificity at
 other positions.

!-: CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
 methionyl peptide.

!-: COFACTOR: Binds 1 iron(II) ion (BY SIMILARITY).

CC BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

DR AF01285; AAC24930.1; -.
 DR HSSP; P27251; DEF.
 DR SubjList; BG11815; defB.
 DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01127; Pep deformylase; 1.
 DR ProDom; PD003444; Pep deformylase; 1.
 DR TIGRFAMS; TIGR00079; Pep deformyl; 1.
 KW Protein biosynthesis; Hydrolase; Iron.
 FT METAL 110 110 IRON (BY SIMILARITY).
 FT METAL 153 153 IRON (BY SIMILARITY).
 ACT SITE 154 154 BY SIMILARITY.
 FT METAL 157 157 IRON (BY SIMILARITY).
 SQ SEQUENCE 184 AA; 20655 MW; 8641BF192666GC38 CRC64;

Query Match 51.8%; Score 575.5; DB 1; Length 184;
 Best Local Similarity 59.5%; Pred. No. 1.7e-40;
 Matches 110; Conservative 33; Mismatches 41; Indels 1; Gaps 1;

Qy 25 MLTMKDIDRSHPTQRKAAELPLPKKEKSTLTAMREFLNSQDEIILKRYGIRSGVG 84
 Db 1 MTMENIVDGHPALETEPVLLPPDAEQQLADMEFVNSQPELLAKYKLRPGV 60

Qy 85 LAAPQINISKRMIAVLLPDGSGSKSYDMLYNPKIVSHSVQRAYLPGEGLSYDDNVAQ 144
 Db 61 LAAPQINISKRMIAV-HADASSEKLYSYALPNPKIVSHSVKVEYSLTSGEGLSYDEA1PG 119

Qy 145 LYERHNKTTIKAQDLEGNDIQLRLKGTPATVQHEDIDHNGMVFHDIDKDHPIQPHHTA 204
 Db 120 YPYRAIRYRGTTLEGENDIDRKLGFPATVQHEDIDHNGMVFHDIDKPNPKREPENA 179

Qy 205 VEYHQ 209
 Db 180 IATR 184

RESULT 2

DEF2_BACST STANDARD; PRT; 184 AA.
 ID_DEF2_BACST
 AC Q31101;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Peptidyl deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase 2).
 DS OS *Bacillus stearothermophilus*.
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.
 NCBI_TaxID=1422; [1]

RESULT 3

DEF2_BACHD STANDARD; PRT; 182 AA.
 ID_DEF2_BACHD
 AC Q9K919;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Peptidyl deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
 GN DBF OR BH2658.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 RN NCBI_TaxID=86665; [1]

SEQUENCE FROM N.A.
 STRAIN=C-125 / JCM 9153;
 MEDLINE=201521582; PubMed=11058132;
 TAKAMI H., Nakasone K., Takagi Y., Maeno G., Sasaki R., Masui N.,
 FUJI F., Hirama C., Nakamura Y., Ogasawara N., Kuwara S.,
 NORIKI K.;
 "Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic comparison with *Bacillus subtilis*."
 Nucleic Acids Res. 28:4317-4331(2000).
 -!- FUNCTION: Removes the formyl group from the N-terminal Met of newly synthesized proteins. Requires at least a dipeptidase for an efficient rate of reaction. N-terminal L-methionine is a prerequisite for activity but the enzyme has broad specificity at other positions (By similarity).
 -!- CATALYTIC ACTIVITY: Formyl-L-methionyl Peptide + H(2)O = formate + methionyl peptide.
 -!- COFACTOR: Binds 1 iron (II) ion (By similarity).
 -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AP001516; BAB0377.1; -.
 DR FSSP; P27251; 2DFE.
 DR InterPro; IPR000181; Pep_deformylase.
 DR Pfam; PF01327; Pep_deformylase; 1.
 DR ProDom; PD003844; Pep_deformylase; 1.
 DR TIGRFAMS; TIGR00079; Pep_deformyl; 1.
 KW protein biosynthesis; Hydrolase; Iron; Complete proteome.
 FT METAL 110 110 IRON (BY SIMILARITY).
 PT ACT SITE 154 154 IRON (BY SIMILARITY).
 PT METAL 157 157 IRON (BY SIMILARITY).
 SQ SEQUENCE 182 AA; 20599 MW; 665D39B56EE6153F CRC64;

Query Match 45.5%; Score 506.5%; DB 1; Length 182;
 Best Local Similarity 58.0%; Pred. No. 7, 9e-35; Mismatches 39; Indels 1; Gaps 1;
 Matches 98; Conservative 31; MisMatches 39; Indels 1; Gaps 1;

Qy 25 MLTMKDIIRDGHPTLROKAELLEPLTKREKETLAMRREFLVNSQDEBEIAKRYGSRSGVG 84
 Db 1 MLTMKDIIVREGNVLREVARPKPVPLSDDKTAKMLPLETLNSQNPIEAHKYSRPGVG 60

Qy 85 LAAPQINISKMAVIALIPDGGSKEYSDWMLNPKVSHSYQEAYLPFGEGLSVTDNVAG 144
 Db 61 LAAPQIGSKMIAVTTDNE-KEYSLVLFPKTISESTYMLLEGEGCLSLREVQG 119

Qy 145 IVERHNKIKITKARDIEGNDIQOLLRKGPAVFOHEIDHNGVMYDHID 193
 Db 120 IVPRHARITVAINEENNEEVRULKGFPAIVFQHEIDHNGIMYDRIE 168

RESULT 4
 DEF _STRN STANDARD; PRT; 203 A.
 AC Q9F2F0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Peptide deformylase (EC 3.5.1.88) (PDB) (Polypeptide deformylase).
 GN DEF OR SP1456.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 QX NCBI_TAXID=1313;
 RP SEQUENCE FROM N.A.
 RN MEDLINE=2115652; PubMed=11257016;
 RX

DD	61	LAANQGLKQKTYAVLIPNPPEVEDGEAELPKYKPAIVQHIDHLNGMFDHYI	192
	-DDFL-LACIA	STANDARD;	PRT; 196 AA.
	DEF LACIA		
	Q8CIO8; Q86611;		
	30-MAY-2000 (Rel. 39, Created)		
	DT 16-OCT-2001 (Rel. 40, Last sequence update)		
	DB 15-JUN-2002 (Rel. 41, Last annotation update)		
DDE	Peptide deformylase (BC 3.5.1.88) (PDB) (Polypeptide deformylase).		
DEF OR LI0560.			
GN	Lactococcus lactis (subsp. lactis) (Streptococcus lactis), Lactococcus.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococaceae; Lactococcus.		
OX	NCBI_TaxID=1360;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RP	SEQUENCE FROM N.A.		
RR	STRAIN=LMM230;		
RC	MEDLINE=95291425; PubMed=773380;		
RC	Cancilla M.R., Hillier A.J., Davidson B.E.;		
RC	"Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene, gap: further evidence for strongly biased codon usage in glycolytic pathway genes." / Microbiology 141:1027-1036(1995).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RR	SEQUENCE FROM N.A.		
RC	STRAIN=LIL403;		
RC	MEDLINE=21235186; PubMed=11337471;		
RA	Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,		
RA	Weissenbach J., Ehrlrich S.D., Sorokin A.;		
RA	"The complete genome sequence of the lactic acid bacterium Lactococcus lactis spp. lactis IL1403." / Genome Res. 11:731-753 (2001).		
CC	-!- FUNCTION: Removes the formyl group from the N-terminal Met of newly synthesized proteins. Requires at least a dipeptide for an efficient rate of reaction. N-terminal L-methionine is a prerequisite for activity, but the enzyme has broad specificity at other positions (By similarity).		
CC	-!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate + methionyl peptide.		
CC	-!- COFACTOR: Binds 1 iron(II) ion (By similarity).		
CC	-!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
DR	EMBL: AAC41454_1; ALT INIT.		
DR	EMBL: AE000290; AAC04658_1; ALT INIT.		
DR	InterPro: IPR00183; Pep_deformylase.		
DR	Pfam: PF01327; Pep_deformylase_1.		
DR	ProDom: PPD03844; Pep_deformylase_1.		
DR	TIGR40073; Pep_deformylase_1.		
DR	Protein biosynthesis; Hydrolase; Iron; Complete proteome.		
FT	METAL_123_123	IRON (BY SIMILARITY).	
FT	METAL_166_166	IRON (BY SIMILARITY).	
FT	ACT SITE_167_167	BY SIMILARITY.	
FT	METAL_170_170	IRON (BY SIMILARITY).	
FT	CONFLICT_86_86	D > E (IN REF. 1).	
SQ	SEQUENCE_196_AA; 22047 MW; C5914BB3A92BF243 CRG64;	2592 MW; 361F434048B505DF CRC64;	
Query Match	38.3%; Score 426; DB 1; Length 196;	20.7%; Score 230; DB 1; Length 216;	
Best Local Similarity	49.2%; Pred. No. 3.5e-28;	30.4%; Pred. No. 4.8e-12;	
Matches	91; Conservative 36; Mismatches 46; Indels 12; Gaps 3;	Matches 59; Conservative 44; Mismatches 83; Indels 8; Gaps 5;	
Qy	25 MLTKMDITRDGRPTLRQXAEELPLTKKEETLJAMREFLNSDDEETAKYGLRSVG 84	7 NIFTVIQ-LKQIKIRKIQYMLTMKDIIRGHPTLRQKAABELPLTKKEETLJAMREFL 65	
Db	1 MISMDDIRCEGPTLEPLREVANDVTPLSDDEDILSERMLQFLNSDQPMWAERGLRGVG 60	11 SFRIPLILPQI---NNMELPTEKAWLVDVKINEETKPYQFPDQASLDCTAKMAYV 67	
Qy	85 LARQINISKRMIAVLP---DGCS---DGCS---GKSYDY---MLVNPIVSHSVQ 132	66 VNSQDEBIAKRYCERSYGLAAPQINISKRMIAVLPIDGGSGKSYDYMVNPKVSHSVQ 125	

QY	119	TYSHSVQOBAYLPTGCGLSYDDNVAGVYHRENKTTIKAD-LEGNDIQQLRKGYPAAVFO	177	35	GHPFLRKQAKAABE-----DLTKEKEKETTLIAMREFLVNSDDEIATKRYGLRSQGVGLAAPQ	89
Db	118	ITDQSENKAFLFESGGCGLSVKKHQKTVIRSEWITIKGCDWPEKKEITIKATGLFGMCIQ	177	9	GDPMLRRKARPYVEDESGIRLAEDMLETMFPAK-----GVGLAAPQ	49
Qy	178	HEIDDHNGVMFYDHIDHDHBLQP	200	90	INISKRM-LAVLIPDGSGSYD-----YMLVNPKIVSHSVQEAYLPGEGLSLVD	139
Db	178	HEPDHLQGRPFYQRI--NPLNP	197	50	IGLSQRLEVATEADPPEGEEPRPLRELVRYVANPVT---TYRBVLVEGTGELSLP	106
RESULT 9				140	DNYAGLYVRHENKTTIKAD-LEGNDIQQLRKGYPAAVFOHEIDHLNGYMFYDHIDK	194
ID - DEF THETH	DEF THETH	STANDARD;	PRT;	107	GLYSSEVERAERIRVEYQDEGRVLELEGYMARVEQHEIDHLGDTLFFERLKP	161
AC P43522;				107	GLYSSEVERAERIRVEYQDEGRVLELEGYMARVEQHEIDHLGDTLFFERLKP	161
DT 01-NOV-1995	(Rel. 32, Created)					
DT 01-NOV-1995	(Rel. 32, Last sequence update)					
DT 15-JUN-2002	(Rel. 41, Last annotation update)					
DEF.						
OS Thermus thermophilus						
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;						
OC Thermaceae; Thermus.						
OX NCBI_TaxID=274;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC STRAIN=YKL;						
RX MEDLINE=950326; PubMed=7961514;						
RA Meirnei T.; Blanquet S.;						
RT "Characterization of the <i>Thermus thermophilus</i> locus encoding peptide deformylase and methionyl-tRNA (fMet) formyltransferase.";						
RT involving three conserved motifs and a metal ion.";						
RL J. Bacteriol. 176:7387-7390(1994).						
RN [2]						
RP CHARACTERIZATION:						
RX MEDLINE=97272005; PubMed=9126850;						
RA Meirnei T.; Lazzennec C.; Villoing S.; Blanquet S.;						
RT "Structure-function relationships within the peptide deformylase family. Evidence for a conserved architecture of the active site involving three conserved motifs and a metal ion.";						
RT J. Mol. Biol. 267:749-761(1997).						
RL J. Mol. Biol. 267:749-761(1997).						
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of newly synthesized proteins. Requires at least a dipeptidase for an efficient rate of reaction. N-terminal L-methionine is a prerequisite for activity but the enzyme has broad specificity at other positions (By similarity).						
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl Peptide + H(2)O = formate +						
CC -!- COFACTOR: Binds 1 iron (II) ion (By similarity).						
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.						
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).						
CC DR X79087; CA55695.1; -.						
CC DR HSSP; P27251; 2DF.						
CC DR InterPro; IPR001081; Pep deformylase.						
CC DR Pfam; PF01327; Pep deformylase; 1.						
CC DR ProDom; PD003844; Pep deformylase; 1.						
CC DR TIGRFAMS; TIGR00079; pep deformyl; 1.						
CC DR TIGRFAMS; TIGR00079; pep deformyl; 1.						
KW protein biosynthesis; Hydrolase; Iron.						
FT METAL 102 102 IRON (BY SIMILARITY).						
FT METAL 145 145 IRON (BY SIMILARITY).						
FT ACT SITE 146 146 BY SIMILARITY.						
FT METAL 149 149 IRON (BY SIMILARITY).						
SQ SEQUENCE 192 AA; 22092 MW; 665941834251361 CRC64;						
Query Match 19-4%; Score 215.5; DB 1; Length 192;						
Best Local Similarity 33.1%; Pred. No. 6.4e-11;						
> Matches 58; Conservative 30; Mismatches 50; Indels 37; Gaps 5;						

5Q SEQUENCE 169 AA; 19057 MW; 7EA00AB510F798CD CRC64;
 Query Match 19.0%; Score 211.5; DB 1; Length 169;
 Best Local Similarity 35.1%; Pred. No. 1.e-10;
 Matches 53; Conservative 31; Mismatches 48; Indels 19; Gaps 6;

Qy 49 PLTKEKEETLIAEREFVNSDDEELAKYGLRGGVLAAPPOINTSKRATVALLPDDGSK 108
 Db 20 PTKND---AIRKEDDMTYQEK---GIGAAQVQDLQRITIDEKG---K 67

Qy 109 SYDMLVNPKVTSYSEAYLPTGCLSYDDNVAGLYHNRNKKITKARDIEGNDIQLRL 168
 Db 68 QNQFLIPEELA---SEGEGIEGCGSI---PGRALVPRKEKTVRALRDGKEFTLDA 123

Qy 169 KGYPAIVNTQEEFDHLNMGMFHDHKDHQ 199
 Db 124 DGLHACIQHQETDHLNGLFVDLS---PLK 151

RESULT 11
DEFNC_ARATH STANDARD; PRT; 273 AA.

AC Q9LYJ4; Q9LYJ4; PRT; 273 AA.
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DS Peptide deformylase, chloroplast precursor (EC 3.5.1.88) (PDE)
 DS (Polypeptide deformylase).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplante; Streptophyt; Embryophyta; Tracheophyta;
 OC eurosidae II; Magnoliophyta; eudicots; Rosidae;
 OC eurosidae II; Brassicales; Brassicaceae; Arabidopsis.
 RN [1] STRAIN=cv; Columbia; PubMed=11060042;
 RX MEDLINE=20514156; PubMed=11060042;
 RA Gilleo C., Serero M., Boisson B., Meimael T.;
 RT "Identification of eukaryotic peptide deformylases reveals
 universality of N-terminal protein processing mechanisms.";
 RL EMBO J. 19:5916-5929 (2000).
 RN [2]

SEQUENCE FROM N_A.
 RN [1] STRAIN=cv; Columbia; PubMed=21016721; PubMed=11130714;
 RX Tabata S., Kaneko T., Nakamura Y., Kojiri T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouki T., Kawashima K.,
 RA Kohara M., Matsutomo M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shimpou S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoenkeling T., Pepin K., Spielet J., Sahron M., Armstrong J., Becker M.,
 RA Belter B., Cordon H., Cordes M., Courtney L., Courtney W., Daitte M.,
 RA Edwards J., Fryman J., Haafenszen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,
 RA Wagner-McPherson C., Wollam A., Yokaum M., Bell M., Dediha N.,
 RA Parneil L., Shah R., Rodriguez M., Hoon L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Barret A., Miller B., Marra M.,
 RA Martianssen R., McCombie W.R., Wilson R.K., Murphy G., Baurodt I.,
 RA Volckert G., Wambutt R., Duesterhoeft A., Stielkema W., Pohl T.,
 RA Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Gromponez B., Zimmermann W.,
 RA Ramsperger U., Wedler H., Wedler E., Peters S.,
 RA van Steveren M., Dirksen W., Mooijman P., Klein Lankhorst R.,
 RA Weitzneger T., Botho G., Rose M., Hauf J., Bernsiger S., Hempel S.,
 RA Feldpusch M., Lamberth S., Villarroel R., Gilien J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rind S., Schoof H.,
 RA Schueler C., Zaccaria P., Meves H.-W., Bevan M., Fausz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana"; Nature 408:823-826 (2000).
 RL FUNCTION: Removes the formyl group from the N-terminal Met of

newly synthesized proteins (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate + methionyl peptide.
 CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SUBCELLULAR LOCATION: Chloroplast (Potential).
 CC -!- SIMILARITY: BELONGS TO THB POLYPEPTIDE DEFORMYLASE FAMILY.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC ---
 DR EMBL; AF262165; AAC33980.1; -;
 DR EMBL; AL163792; CAB87633.1; -;
 DR HSSP; P27251; 1DER;
 DR InterPro; IPR00181; Pep deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR ProDom; P00384; Pep deformylase; 1.
 DR TIGRFAM; TIGR00079; Pep deformylase; 1.
 KW CHLOROPLAST (POTENTIAL).
 FT TRANSIT 1 ?
 FT CHAIN ?
 FT METAL 171 171
 FT METAL 213 213
 FT ACT SITE 214 214
 FT METAL 217 217
 FT CONFLICT 205 205
 SQ SEQUENCE 273 AA; 30541 MW; D188A90D768D03F CRC64;
 SQ ---
 Query Match Score 18.7%; Score 208; DB 1; Length 273;
 Best Local Similarity 32.2%; Pred. No. 4.1e-10;
 Matches 55; Conservative 33; Mismatches 65; Indels 18; Gaps 6;
 Qy 22 VOYMLTMKDIIRDGHPTLQRKAELPLPKTEKEETLAMREFLYNSQDEELAKRYGRS 81
 Db 75 VQEFETPK-VEYPDPLTRAKNKRDI-EDEENKVLADAMPFVYKT-----D 120
 Qy 82 GYGLAQPQINLQML-VENNAAGEPGEKGKVLVNPKIKYS--DKLVPDPGEGLSPFGI 141
 Db 121 GIGLSPQVQNLVQML-VENNAAGEPGEKGKVLVNPKIKYS--DKLVPDPGEGLSPFGI 177
 Qy 142 VAGLVRHNRKTTIKAKDIEGNDIQLRKGPALYFOEIDHNGMNEYDHL 192
 Db 178 YAEEV-RPQSVKTDARDITGFRFSISSLSPARIFQHEYDHLEGVLFFDRM 227
 RESULT 12
 DEF CHLMU STANDARD; PRT; 181 AA.
 AC QPK41;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Peptide deformylase (EC 3.5.1.88) (PDE) (Polypeptide deformylase).
 RN DEF_TC032.
 OS Chlamydia muridarum.
 OC NCBI_TaxID=3560;
 RN [1] SEQUENCE FROM N_A.
 RP STRAIN=MOpN / Nigg
 RC MEDLINE=20150255; PubMed=10581935,
 RX Read T.D., Brunham R.C., Shen C., Gill S.B., Heidelberg J.F., Bass S.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Knoury H., Craven B., Bowman R., Dobson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 pneumoniae AR3";
 RT Nucleic Acids Res. 28:1397-1406 (2000).

-!- FUNCTION: Removes the formyl group from the N-terminal Met of newly synthesized proteins. Requires at least a dipeptide for an efficient rate of reaction. N-terminal L-methionine is a prerequisite for activity but the enzyme has broad specificity at other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate + methionyl peptide.
 CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.

CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR EMBL; AE001308; AAC67948.1; -.
 CC DR HSSP; P27251; 2DBF.
 CC DR InterPro; IPR00181; Pep_deformylase.
 CC DR TIGER; TC0632; -.
 CC DR InterPro; IPR00181; Pep_deformylase; 1.
 CC DR PFam; PF01327; Pep_deformylase; 1.
 CC DR PRODOM; PD003844; Pep_deformylase; 1.
 CC DR TIGRFAMs; TIGR00079; Pep_deformyl; 1;
 CC KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
 FT METAL 99 99
 FT METAL 141 141
 FT ACT SITE 142 142
 FT METAL 145 145
 SQ SEQUENCE 181 AA; 20523 MW; A3FD251E4DFB96DA CRG64;
 SQ SEQUENCE 181 AA; 20535 MW; 5EB2DD7C0A3395CA CRC64;

Query Match 18.3%; Score 203; DB 1; Length 181;
 Best Local Similarity 33.5%; Pred. No. 6.4e-10;
 Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;
 Qy 31 IIRD---GIFTILOKKAABELPLTKEKEKETLIAMREFTVNSQDEELAKRYGLRSGVLA 86
 Db 1 MIRDLEYDTSILLRKVAAPV-TBTITBLRLVLMSETM-----AFYKGIVGLA 47
 Qy 87 APQINISKEMIAILIP--DDGSGKSYD--MLYNPKTVSHSQEAYLPFGEGCLSVDDN 141
 Db 48 APQVGHSAVALIMGYKEELDGGELIFDFPKVTFNPVTKTQSEQLY--GNNEGLSI-PG 104
 Qy 142 VAGLVERHNKTTIKDLEGNDIQLLRKGYPATVQEHDIDHGMVYDH-IDDHPLQ 129
 Db 105 LREBEVARPDKITVSQKNDQQESLAEGFLGARYWMDHHEVLYIDMSDDDKTKQ 163

RESULT 14
 DEF _THEMA
 ID _DEF THEM
 AC P96713
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
 OC Thermotogaceae; Thermotoga.
 OC NCBI_TaxID=2336;
 RN [1] RP SEQUENCE FROM N.A.
 RP STRAIN=DTUW-3/Cx;
 RX MEDLINE=99000809; PubMed=9784136;
 RA Stephens R.S.; Kalman S.; Lammel C.J.; Fan J.; Marathe R.; Aravind L.; Mitchell W.P.; Olinger L.; Tatusov R.L.; Zhao Q.; Koonin E.V.; Davis R.W.; RT "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";
 RT Science 282:754-759 (1998).
 RL -!- FUNCTION: Removes the formyl group from the N-terminal Met of newly synthesized proteins. Requires at least a dipeptide for an

CC efficient rate of reaction. N-terminal L-methionine is a prerequisite for activity but the enzyme has broad specificity at other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate + methionyl peptide.
 CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR EMBL; AE001308; AAC67948.1; -.
 CC DR PHCPI-2DPIAGE; 084357; -.
 CC DR InterPro; IPR000181; Pep_deformylase.
 CC DR Pfam; PF01327; Pep_deformylase; 1.
 CC DR PRODOM; PD003844; Pep_deformylase; 1.
 CC DR TIGRFAMs; TIGR00079; Pep_deformyl; 1;
 CC KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
 FT METAL 99 99
 FT METAL 141 141
 FT ACT SITE 142 142
 FT METAL 145 145
 SQ SEQUENCE 181 AA; 20523 MW; A3FD251E4DFB96DA CRG64;
 SQ SEQUENCE 181 AA; 20535 MW; 5EB2DD7C0A3395CA CRC64;

Query Match 18.3%; Score 203; DB 1; Length 181;
 Best Local Similarity 33.5%; Pred. No. 6.4e-10;
 Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;
 Qy 31 IIRD---GIFTILOKKAABELPLTKEKEKETLIAMREFTVNSQDEELAKRYGLRSGVLA 86
 Db 1 MIRDLEYDTSILLRKVAAPV-TBTITBLRLVLMSETM-----AFYKGIVGLA 47
 Qy 87 APQINISKEMIAILIP--DDGSGKSYD--MLYNPKTVSHSQEAYLPFGEGCLSVDDN 141
 Db 48 APQVGHSAVALIMGYKEELDGGELIFDFPKVTFNPVTKTQSEQLY--GNNEGLSI-PG 104
 Qy 142 VAGLVERHNKTTIKDLEGNDIQLLRKGYPATVQEHDIDHGMVYDH-IDDHPLQ 129
 Db 105 LREBEVARPDKITVSQKNDQQESLAEGFLGARYWMDHHEVLYIDMSDDDKTKQ 163

RESULT 14
 DEF _THEMA
 ID _DEF THEM
 AC P96713
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
 OC Thermotogaceae; Thermotoga.
 OC NCBI_TaxID=2336;
 RN [1] RP SEQUENCE FROM N.A.
 RX MEDLINE=97240832; PubMed=9086272;
 RA Mazel D.; Coic E.; Blanchard S.; Saurin W.; Marliere P.; RT "A survey of polypeptide deformylase function throughout the eubacterial lineage"; RL J. Mol. Biol. 266:939-949 (1997).
 RA [2] RP SEQUENCE FROM N.A.
 RX STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=9287316; PubMed=10360571;
 RA Nelson K.E.; Clayton R.A.; Gill S.R.; Gwinn M.L.; Dodson R.J.; Haft D.H.; Hickey E.K.; Peterson J.D.; Nelson W.C.; Ketchum K.A.,

RESULT 15						
	DEF_BUGAL	STANDARD;	PRT;	173 A.		
Qy	ID P57563;					
Db	DT 16-OCT-2001 (Rel. 40, Created)					
Qy	DT 16-OCT-2001 (Rel. 40, Last sequence update)					
Db	DT 15-JUN-2002 (Rel. 41, Last annotation update)					
Qy	DN Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).					
Db	DN DEF OR BU490.					
Qy	GN Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium)					
Db	OS Bacterium; Proteobacteria; Gamma subdivision; Buchnera.					
RN [1] _	NCBI_TaxID=118099;					
RN [1] _	SEQUENCE FROM N.A.					
RP STRAIN=TOKYO 1988.						
RC						

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw mode¹

Run on: July 11, 2003, 09:31:30 ; Search time 80 Seconds
(without alignments)
548.600 Million cell updates/sec

Title: US-09-896-580A-1
Perfect score: 1112
Sequence: 1 MYEYLNLLFTVIQKQIKIR..... KDHPLQPHTDAVEHQHHHH 213

Scoring table: BLOSUM62
GapOp 10.0 , GapExt 0.5

Searched: 671580 seqs, 20047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : SPREMBL 21:

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	% Match	Length	DB ID	Description
1	938	84.4	183	16	Q9F14	Q9F14 staphylococ
2	536.5	48.2	183	16	O92CX8	O92cx8 listeria in
3	533.5	48.0	183	16	Q8Y866	Q8Y866 listeria mo
4	477	42.9	204	16	Q9XXY7	Q9XXY7 streptococc
5	462.5	41.6	203	2	Q93R9	Q93R9 streptococc
6	274.5	24.7	204	2	Q9CK63	Q9CK63 mycoplasma
7	273.5	24.6	198	16	Q9PQ25	Q9PQ25 ureplasma
8	263	23.7	198	16	Q8PN3	Q8PN3 mycoplasma
9	226.5	20.4	224	16	Q92J7	Q92J7 rickettsia
10	218.5	19.6	168	16	Q9XN16	Q9XN16 vibrio chol
11	217.5	19.6	179	16	Q9YVH1	Q9YVH1 anabaena sp
12	215	19.3	178	2	Q93IE9	Q93IE9 leptospira
13	206.5	18.6	178	16	Q9ZSH6	Q9ZSH6 rhizobium m
14	206	18.5	273	10	Q94U8	Q94U8 arabidopsis
15	205.5	18.5	232	16	Q8RRO4	Q8RRO4 deinococcus
16	199	17.9	147	16	Q8XJL2	Q8XJL2 clostridium

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1	ID	Q9F14	PRELIMINARY;	PRT;	183 AA.
	AC	Q9FL4;			
	DT	01-MAR-2001 (T-EMBL; 16, Created)			
	DT	01-MAR-2001 (T-EMBL; 16, Last sequence update)			
	DT	01-JUN-2002 (T-EMBL; 21, Last annotation update)			
	DE	Peptide deformylase Pdfl (Pdfl protein) (Formylmethionine deformylase homolog).			
	DN	DEFI OR Pdfl OR SAV1091 OR SA0442.			
	OS	Staphylococcus aureus (strain Mu50 / ATCC 700699),			
	OS	Staphylococcus aureus (strain N315), and			
	OS	Staphylococcus aureus.			
	OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;			
	OC	Staphylococcus.			
	NCBI_TaxID	58878, 158879, 1280;			
	RN	[1]			
	RP	SEQUENCE FROM N.A.			
	RC	SPECIES-S. aureus; STRAIN=WCUH29;			
	RA	Lonetto M.A., Sylvester D.R., Warren R.L.;			
	RT	"Staphylococcus aureus deformylase 1 encoding DNA."			
	RL	Submitted (AUG-2000) to the ENB/GenBank/DBJ databases.			
	RN	[2]			
	RP	SEQUENCE FROM N.A.			
	RC	SPECIES-S. aureus; (strain Mu50), and S. aureus (strain N315); MEDLINE=1311952; PubMed=11418146;			
	RX	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Onchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T., Kanamori M., Matsunari H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Uji Y., Takanashi N.K., Inoue S., Yabuzaiki J., Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaiki J., Kaneshisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hirunatsu K.,			
	RA	"Whole genome sequencing of methicillin-resistant Staphylococcus aureus".			
	RA	Lancet 357:1225-1240 (2001).			
	DR	EMBL; AY007227; ARGO229.1; -			
	DR	EMBL; AP003361; BAB5723.1; -			
	DR	EMBL; AP003132; BAB42188.1; -			

DR	HSSP; P27251; 1DEF.	Qy	25 MLLTMKDLIRGHPHTLPRQKAELPLTKEEKTLLIAMR3FLVNSQDEIARKGRLRSVG 8-4
DR	InterPro; IPR000181; Pep deformylase.	Db	1 MLTMDDIVRECHPALETVETFPLSDEEKLGHDMLFLINSDQDELAERKGRLRGCGV 6-0
DR	Pfam; PF01577; Pep deformylase.	Qy	85 LAAPQINISKRMIAVLIPDGGSKSYDYMVNPKVSHSVQAYLPTEGCLSYDDNVAG 14-4
DR	PRINTS; PRO1576; PDEFORMYLASE.	Db	61 TARPQLAVTRPAFHVADE-KDRLKSYVLMNPKRSVHQACLSGEGCLSYDREVPG 11-9
DR	ProDom; PDO03844; Pep deformylase; 1.	Qy	145 LYERHNKIKTAKDIEGNDIQLRLKGYPATVQEHIDHNGMFYDHIDKDP--LQPHT 20-2
DR	TIGRFAMS; TIGR00079; pep_deformyl; 1.	Db	120 YYVRSRVTIDADEFNGTPKLRFKGYPATVQEHIDHNGMFYDHINKENPSYLPDPV 17-9
KW	Complete Proteome.	SEQUENCE	183 AA; 20558 MW;
SEQUENCE	32A64066A6FAABOE CRC64;	Qy	25 MLTMKDLIRGHPHTLPRQKAELPLTKEEKTLLIAMR3FLVNSQDEIARKGRLRSVG 8-4
Query Match	84.4%; Score 938; DB 16; Length 183;	Db	1 MLTMKDLIRGHPHTLPRQKAELPLTKEEKTLLIAMR3FLVNSQDEIARKGRLRSVG 6-0
Best Local Similarity	98.9%; Pred. No. 1_5e-73;	Matches	181; Conservative
Matches	2; Mismatches 0; Gaps 0;	Indels	0;
Db	61 LAAPQINISKRMIAVLIPDGGSKSYDYMVNPKVSHSVQAYLPTEGCLSYDDNVAG 14-4	RESULT	3
Qy	85 LAAPQINISKRMIAVLIPDGGSKSYDYMVNPKVSHSVQAYLPTEGCLSYDDNVAG 84	Qy	85 QY8566 PRELIMINARY; PRT; 183 AA.
Db	61 LAAPQINISKRMIAVLIPDGGSKSYDYMVNPKVSHSVQAYLPTEGCLSYDDNVAG 120	Db	Q8YX66; DT 01-MAR-2002 (TREMBrel. 20, Created)
Qy	145 LYERHNKIKTAKDIEGNDIQLRLKGYPATVQEHIDHNGMFYDHIDKDP--LQPHT 204	Db	DT 01-MAR-2002 (TREMBrel. 20, Last sequence update)
Db	121 LYERHNKIKTAKDIEGNDIQLRLKGYPATVQEHIDHNGMFYDHIDKDP--LQPHT 180	Qy	205 VEV 207
Qy	205 VEV 207	Db	DT 01-JUN-2002 (TREMBrel. 21, Last annotation update)
Db	181 VEV 183	Qy	180 D 180
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	Db	DE Hypothetical protein lmo1051.
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	GN	lmo1051.
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	RN	NCBI_TaxID=1639;
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	RN	SEQUENCE FROM N.A.
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	AC	RP STRAIN=BGD-E / SEROVAR 1/2A;
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	AC	RC MEDLINE=2157275; PubMed=11679659;
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Charbit A., Chetouani F., Couve E., de Darvuar A., Dehoux P.,
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurgeot O.,
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Jones L.-M., Kaerst U., Kraft J., Kuhn M., Kunst F., Kurapkat G.,
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Madueno B., Maitournam A., Mata Vicente J., NG E., Nedjari H.,
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Remmel B., Rose M., Schlüter T., Simoes N., Tierrez A.,
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DT	RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	RT Comparative genomics of Listeria species.";
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	Science 294:849-852(2001).
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	EMBL AL591977; CAS39139.1; -.
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	ListList; LMO01051; -.
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	InterPro; IPRO00181; Pep deformylase.
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	PFam; PF01327; Pep deformylase; 1.
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	PRINTS; PRO1576; PDEFORMYLASE.
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	ProDom; PDO03844; Pep deformylase; 1.
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	TIGRFAMS; TIGR00079; Pep deformyl; 1.
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	Hypothetical protein; Complete proteome.
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	SO	SEQUENCE 183 AA; 20543 MW; 65B2430603CD4EF CRC64;
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	SO	Query Match 48.0%; Score 533.5; DB 16; Length 183;
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	SO	Best Local Similarity 57.5%; Pred. No. 1_7e-38; Mismatches 44; Indels 3; Gaps 2;
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	25 MLTMDDIVRECHPALETVETFPLSDEEKLGHDMLFLINSDQDELAERKGRLRGCGV 84
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	DR	1 MLTMDDIVRECHPALETVETFPLSDEEKLGHDMLFLINSDQDELAERKGRLRGCGV 60
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	Qy	85 LAAPQINISKRMIAVLIPDGGSKSYDYMVNPKVSHSVQAYLPTEGCLSYDDNVAG 14-4
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	Db	61 TARPQLAVTRPAFHVADE-KDRLKSYVLMNPKRSVHQACLSGEGCLSYDREVPG 11-9
Db	Q92CX8 PRELIMINARY; PRT; 183 AA.	Qy	145 LYERHNKIKTAKDIEGNDIQLRLKGYPATVQEHIDHNGMFYDHIDKDP--LQPHT 20-2
Qy	Q92CX8 PRELIMINARY; PRT; 183 AA.	Db	120 YYVRSRVTIDADEFNGTPKLRFKGYPATVQEHIDHNGMFYDHINKENPSYLPDPV 17-9

DR	InterPro; IPR000181; Pep_deformylase.	QY	150 NKITIKAKD-IEGNDIQLRILKGYPAIVFQHEIDHLNGMFYDHIDKDHF 197
DR	PFam; PF03227; Pep_deformylase.	DB	138 KKVHVTAYDILSEQFDQDFSGIAICIOEIGHLDAGYYTNINQQP 186
DR	PRINTS; PRO1576; Pep_deformylase.		
DR	ProDom; P00344; Pep_deformylase.		
DR	TIGRFAMS; TIGR0079; Pep_deformylase.		
KW	Hydrolyase; Protein biosynthesis; Zinc.		
SEQUENCE	204 AA; BB5584F1123B7B44 CRC64;		
SQ	23878 MW;		
RESULT 8			
ID	Q98PN3	PRT;	PRELIMINARY; 198 AA.
AC	Q98PN3;		
DT	01-OCT-2001 (TREMBrel. 18, Created)		
DT	01-OCT-2001 (TREMBrel. 18, Last sequence update)		
DT	01-JUN-2002 (TREMBrel. 21, Last annotation update)		
DE	Polypeptide deformylase 2 (PDF2) (Formylmethionine deformylase 2)		
DE	(EC 3.5.1.31).		
MW	680.		
OS	Mycoplasma pulmonis.		
OC	Bacteria; Firmicutes; Clostridium group; Mollicutes;		
OC	Mycoplasmataceae; Mycoplasma.		
NCBI_TaxID	107.		
RN	[1]		
SEQID	FROM N.A.		
RC	STRAIN=DAB CTIP;		
RP	PubMed=11353084;		
RX	MEDLINE=21267165;		
RA	Heiling R., Perris S., Barbee V., Samson D., Galisson F.,		
RA	Chambaud I., Heiling R., Perris S., Barbee V., Samson D., Galisson F.,		
RA	Moszer I., Dublewski H., Wroblewski H., Viari A., Rocha E.P.C.,		
RA	Blanchard A.;		
RT	"The complete genome sequence of the murine respiratory pathogen		
RT	Mycoplasma pulmonis."		
RL	Nucleic Acids Res. 29:2145-2153 (2001).		
DR	EMBL; AL445565; CAC13862.1;		
DR	MyPlis; MWP0.6890; -.		
DR	InterPro; IPR000181; Pep_deformylase.		
DR	Pfam; PF01127; Pep_deformylase; 1.		
DR	PRINTS; PRO1576; PDFORMYLASE.		
DR	ProDom; PD00344; Pep_deformylase; 1.		
DR	TIGRFAMS; TIGR00079; Pep_deformyl; 1.		
KW	Hydrolase; Complete proteome.		
SQ	SEQUENCE 198 AA; 23040 MW; 1A36853577F91CEA CRR64;		
Query Match 274.5; DB 2; Length 204;			
Best Local Similarity	32.1%; Pred. No. 5.4e-16;		
Matches	63; Conservative 45; Mismatches 85; Indels 3; Gaps 3;		
QY	15 KQIKRKVQYMLTMKDIRDGHPTLRQKAELPLPKTEKEETLAMREFLYNSQDEIA 74		
DB	4 KNIFIMKSKRKPTRDMLTDPRKPEEVSEVKPLSEVLDLIDNLAVVDSEFDIN-A 62		
QY	75 KRYGLRSVGEVLAQFOINNSKRMIAVLIPPDGSCKSYDMVLPKTVSHSYOEAFLPFGEG 134		
DB	63 BKYDPRGIGIAANOLGINQNQRFYVFH-LDFCQEKRHLINPEWIDKSLNKAYLAGEG 121		
QY	135 CLSVDDNVAQCLVIRHNLTIKAKD-IEGNDIQLRILKGYPAIVFQHEIDHLNGMFHID 193		
DB	122 CLSVPKDKOGYVSETKLKSFEDVLTQDVTELSAHHGLANCOLHEMDHLEGRFFYDSIN 181		
QY	194 RDHPDQPHDAVEVHQ 209		
DB	182 MMKPKHKGKDEWVCSEQ 197		
RESULT 7			
ID	Q9PQ25	PRT;	PRELIMINARY; 198 AA.
AC	O9PQ25;		
DT	01-OCT-2000 (TREMBrel. 15, Created)		
DT	01-OCT-2000 (TREMBrel. 15, Last sequence update)		
DT	01-JUN-2002 (TREMBrel. 21, Last annotation update)		
DE	Polypeptide deformylase.		
GN	DEF OR UT465.		
OS	Ureaplasma parvum (Ureaplasma urealyticum biotype 1)		
OC	Bacteria; Firmicutes; Clostridium group; Mollicutes;		
OC	Mycoplasmataceae; Ureaplasma.		
OX	NCBI_TaxID=134821;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SEROVAR 3.		
RX	MEDLINE=200500219; PubMed=11048724;		
RA	Glass J.I., Lefkowitz B.J., Glass J.S., Heiner C.R., Chen E.Y.,		
RA	Cassell G.H.;		
RT	"The complete sequence of the mucosal pathogen Ureaplasma		
RT	urealyticum."		
RT	Nature 407:757-762 (2000).		
DR	EMBL; AE002143; AAJ30877.1; -.		
DR	HSSP; P27251; 2DEP.		
DR	InterPro; IPR000181; Pep_deformylase.		
DR	ProDom; P01327; Pep_deformylase; 1.		
DR	PRINTS; PRO1576; PDFORMYLASE.		
DR	ProDom; PD00344; Pep_deformyl; 1.		
DR	TIGRFAMS; TIGR00079; Pep_deformyl; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 198 AA; 23143 MW; 19F035D02625C95 CRC64;		
RESULT 9			
ID	Q92UJ7	PRT;	PRELIMINARY; 224 AA.
AC	Q92UJ7;		
DT	01-DEC-2001 (TREMBrel. 19, Created)		
DT	01-DEC-2001 (TREMBrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBrel. 21, Last annotation update)		
DE	DEP3 OR RC0080.		
GN	Rickettsia conorii.		
OS	Rickettsia; Proteobacteria; alpha subdivision; Rickettsiales;		
OC	Rickettsiaceae; Rickettsiae; Rickettsia.		
NCBI_TaxID	781;		

RN	[1]	SQ	SEQUENCE	168 AA;	18671 MW;	F51EDDE3B6C14C59 CRC64;
RC	SEQUENCE FROM N.A.					
STRATN=MAJISH 7;						
RX MEDLINE=21442074; PubMed=11557893;						
OGATA H., AUDIC S., RENESTO-AUDIFFREN P., FOURNIER P.-E., BARBE V., SAMSON D., ROUX V., COSSART P., WEISSENBACH J., CLAVERIE J.-M., RAOUlt D.; "Mechanisms of evolution in rickettsia conorii and R. prowazekii."; RIL Science 293:2093-2098(2001).						
EMBL; AB008575; AU02618.1; -.						
InterPro; IPR000151; Pep deformylase.						
PFam; PF01327; Pep deformylase; 1.						
PRINTS; PRO1576; PDEFORMYLASE.						
DR Dom; PP00384; Pep deformylase; 1.						
KW Complete proteome.						
SEQUENCE 224 AA;	25539 MW;	74ED5C861C16237F CRC64;				
Query Match	20.4%	Score 226.5;	DB 16;	Length 224;		
Best Local Similarity	34.1%;	Pred. No. 8.8e-12;				
Matches 72; Conservative 29; Mismatches 73; Indels 37; Gaps 9;						
Db 12 IQLQIKRKYQMLTKDIRDGEPHTIQKAALEPLTKE-KETIYAMREFLNSQD 70						RESULT 11
Db 29 ISLPQVTLNQTINNSED-----KTRIKATLNFPLSSEDLDLDISTLKKY--DQE 79						Q8YVHL PRELIMINARY; PRT; 179 AA.
Qy 71 EBIAKRYGLRSVGLQAPQINISLMAVLPDGSKSY-DYMLVNPKVSHSYQEAY 128						AC Q8YVHL
Db 80 ENCA-----GLAAPQIGSKCIIIAVHDAELKKWPHDIDKTMPKTI-WINPSY 128						DT 01-MAR-2002 (TREMBLrel. 20, Created)
Qy 129 LPTG-----EGCLSYDDNTAGLYVRHKITKAKDIEENDIQRLRGYPAVFOHEID 181						DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
Db 129 KPIGDKHEDYEGGPSV-ENATGPVAFKKHYHAYDINGNQIQTARGFLAZVHQHEID 187						DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
Qy 182 HLNGWMMFYHI-----DKDPELIQHTDAVE 206						DE Polypeptide deformylase.
Db 188 HLNGKVFELDYVAPKKKINTKEYLENRKKME 218						GN ALL007.
RESULT 10						OS Anabaena sp. (strain PCC 7120)
OPRN16 ID Q9KNL6 PRELIMINARY; PRT; 168 AA.						OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
AC Q9KNL6 PRELIMINARY; PRT; 168 AA.						RN [1] NCBI_TaxID:103690;
RC STRATN=EL TOR N16561 / SEROTYPE O1;						RP SEQUENCE FROM N.A.
RX MEDLINE=21406833; PubMed=10952301;						RX MEDLINE=21195235; PubMed=11759840;
Heidelberg J.-F., Eisen J.B., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.B., Read T.D., Tetelin H., Richardson D., McDonald L., Utterback T., Blasczmann J., Qin H., Dragoi I., Sellers P., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.; "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.", Nature 406:477-483 (2000).						RA Kaneko T., Nakamura Y., Walk C.P., Kuritz T., Sasamoto S., Watanabe A., Iriuchi M., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M., Yasuda M., Tabata S.; "Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120.", PL DNA Res. 8:205-213 (2001)
VCA0150; VCA0150; Vibriocholera.						PL EMBL; AP003587; BAB3706..1; -.
OC Bacterium; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.						DR InterPro; IPR000181; Pep deformylase.
OX NCBI_TaxID:666;						DR Pfam; PF01327; Pep deformylase; 1.
RN						DR PRINTS; PRO1576; PDEFORMYLASE.
RP SEQUENCE FROM N.A.						DR TIGRFAMS; TIGR00079; _ppt_deformyl1; 1.
RC STRATN=EL TOR N16561 / SEROTYPE O1;						KW Complete proteome.
RX MEDLINE=21406833; PubMed=10952301;						SQ 20136 MW; E17D0BC9B0E473D2 CRC64;
Heidelberg J.-F., Eisen J.B., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.B., Read T.D., Tetelin H., Richardson D., McDonald L., Utterback T., Blasczmann J., Qin H., Dragoi I., Sellers P., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.; "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.", Nature 406:477-483 (2000).						Qy 19.6%; Score 217.5; DB 16; Length 179;
DR EMBL; AE0043556; AAF96063.1; -.						Best Local Similarity 33.1%; Pred. No. 3.9e-11;
DR HSSP; P27251; IDFF.						Matches 57; Conservative 30; Mismatches 60; Indels 25; Gaps 5;
DR InterPro; IPR000181; Pep deformylase.						Qy 31 IIRDGEPTIKRQAKABLEPLTKEETLAMREFLYNSQDEEIAKRYGLRSVGLAIFIQI 90
PFam; PF0327; Pep deformylase; 1.						Db 8 IIQLQPTIKRQAKAWVE-----NHDATIQQLDDLIATVAKANGGIASQV 55
DR PRINTS; PRO1576; PDEFORMYLASE.						Qy 91 NISKRMIAVLPDGSKSY-----MLVNPKYSHSYEAVLPTGRGCJSYDDNA 143
DR Dom; PP003844; Pep deformylase; 1.						Db 56 ASYRPFIVTA--SRPNFYPMEPTAMNPKVGHSTE--IVEGNEGCSV-PGIR 109
TIGRFAMS; TIGR0019; Pep deformylase; 1.						Qy 144 GIYHRKNTIKRQAKDIRENDQLRKGYPALVFEHQHDLANGVMFDHDKD 195
Complete proteome.						Db 110 GLYPRHQATEVEYDTRGNLNLQXQTLTDFTVARIFQHEPDHLDGVLFDVESN 161
KW						RESULT 12

Db 1.21 GIGLSAPQVGLNVQLM-VENPAGEPEGEKIEIVLYNPKKYS--DXLVPFDEGCLSPFGI 177
 Qy 142 VAGLVVERHNKITRAKDFEGNDIQQLKGTPAIVFOHEIDHNGMNFYDH 192
 Db 178 YAEEVV-RPQSVKIDARDITGERPSISLRLPARTIQHEYDLEGYLFDRM 227

RESULT 15

Q9RQ4 PRELIMINARY; FRT; 232 AA.
 AC Q9RQ4; ID 142 VAGLVVERHNKITRAKDFEGNDIQQLKGTPAIVFOHEIDHNGMNFYDH 192
 DT 01-MAY-2000 (TRIMBurrel. 13; Last sequence update)
 DT 01-JUN-2000 (TRIMBurrel. 13; Last annotation update)
 DT 01-JUN-2002 (TRIMBurrel. 21; Last annotation update)
 DE Polypeptide deformylase.
 DR 2434.
 GN OS Deinococcus radiodurans.
 OC Deinococcaceae; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcus.
 OX NCBI_TAXID1299; Deinococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RL;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Moffat K.S., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Vanathavan J.J., Qin H., Jiang J., Pamphilis W., Crosby M., Shan M.,
 RA Makarova K.S., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Ketchum K.A., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Fraser C.M.;
 RA "Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR ENBL; AE002073; AAP11975.1; -.
 DR HSSP; P27251; 2DEF.
 DR TIGR; DR2434;
 DR InterPro; IPR001181; Pep_deformylase.
 DR PF01327; Pep_deformylase.
 DR PRINTS; PRO01576; PFORMYLASE.
 DR PDO3844; PDO deformylase; 1.
 DR TIGRFAMs; TIGR00079; Pep_deformyl; 1.
 DR Complete proteome.
 KW Sequence 232 AA; 26369 MW; 94A31686AE872ED8 CRC64;
 SQ

Query Match Score 205.5; DB 16; Length 232;
 Best Local Similarity 18.5%; Pred. No. 6.1e-10;
 Matches 59; Conservative 34; Mismatches 71; Indels 29; Gaps 7;

Qy 35 GHPTLRQKAEI---ELPLTKEEKSTLJAMREFLVNSQDEIAKRYGLRSQVLAPOI 90
 Db 28 GDPILRKAKNLNTADTLHVPGPFEPQTVREVADMLETMEFER-----RGVGLAAPOI 79

Qy 91 NISKRM-TAVLIPD--GSGES-----YDWMLNP--KIVSESVOEAQYLPTEGSCL 136
 Db 80 GLPVNRVAVAYADDEEENEGQETPLRSVYRETYMLNPKVVKKKDSYY--QBGCL 136

Qy 137 SVDDNTA GLYHRNKRITKAIDCIEGNDIQQLKGTPAIVFOHEIDHNGMNFYDH 196
 Db 137 SPPGIVDGYPRARQVRVDTLDGDPRISEAIDLARYFQHETDLDGLFLDHLPA DI 196

Qy 197 PQQPHDAVEVHQ 209
 Db 197 TEDEHKLRLIRQQ 209